

09/847,513 (MBA-101) Page 4

## STATEMENT TO COMPUTER DISK AND SEQUENCE LISTING

The content of the incorporated sequence listing information recorded in computer readable form is identical to the herein incorporated written sequence listing and no new matter has been included. A written sequence listing of 65 sequences is included as well as a computer disk labeled "Corrected Sequence Listing" for application entitled "Light-driven energy generation using proteorhodopsin" by Edward F. DeLong and Oded Beja" containing files "MBA101-SEQLIST\_CORR.prj", dated "8/4/01" with 171,574 bytes, which is the PatentIn project file generated using PatentIn Version 3.0 software provided by the USPTO, and "MBA101-SEQLIST\_CORR.txt", dated "08/04/01" with 314,695 bytes, which is the generated sequence listing from the PatentIn project file MBA101-SEQLIST\_CORR.prj using PatentIn Version 3.0 software, all which are herein incorporated. The information recorded in computer readable format on the incorporated computer disk labeled "Corrected Sequence Listing" containing files "MBA101-SEQLIST\_CORR.prj" and "MBA101-SEQLIST\_CORR.txt" are identical to the incorporated written sequence listing.

I declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both under Title 18, §1001 of the United States Code, and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Respectfully submitted,

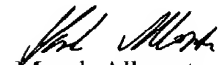


Marek Alboszta  
Reg. No. 39,894

Lumen  
45 Cabot Avenue, Suite 110  
Santa Clara, CA 95051-6670  
(408) 260-7300 x15

following page that the content of the sequence listing recorded in computer readable form on the incorporated computer disk is identical to the incorporated written sequence listing.

Respectfully submitted,

  
Marek Alboszta  
Reg. No. 39,894

Lumen  
45 Cabot Avenue, Suite 110  
Santa Clara, CA 95051-6670  
(408) 260-7300 x15

# SEQUENCE LISTING

<110> MBARI

DeLong, Edward  
Beja, Oded

<120> Light-driven energy generation using proteorhodopsin

<130> MBA-101

<140> US/09/847,513

<141> 2001-05-01

<150> 60/201,602

<151> 2000-05-03

<160> 65

<170> PatentIn version 3.0

<210> 1

<211> 105184

<212> DNA

<213> Naturally occurring gamma proteobacterium

<220>

<221> gene

<222> (50866)..(51615)

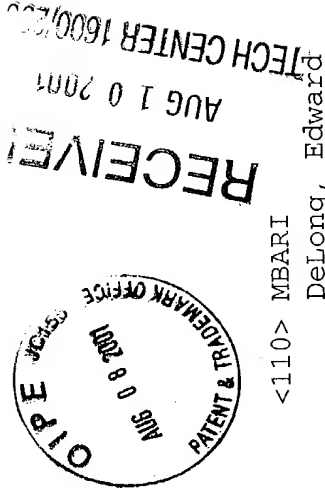
<223> Proteorhodopsin gene sequence.

<220>

<221> misc\_feature

<222> (1593)..(2807)

<223> Predicted threonine dehydratase. Contains 'n' at position 2753.



<300> Beja,O., Aravind,L., Koonin,E.V., Suzuki,M.T., Hadd,A., Nguyen,L.P.,  
 <301> Jovanovich,S.B., Gates,C.M., Feldman,R.A., DeLong,E.F  
 <302> Bacterial rhodopsin: evidence for a new type of phototrophy in the sea  
 <303> Science  
 <304> 289  
 <305> 5486  
 <306> 1902-1906  
 <307> 2000-09-15  
 <308> AF279106  
 <309> 2000-06-15  
 <313> (50866)..(51615)

```

<400> 1                                60
tggttatatc agtaatggct attgctccaa taacttaata ctaatatata attagtttat

gaataaaatt tatatatattg ggttattggt ttttacacta aatgcatttt cttgctcaga 120

tcttctagat acagacatga gagttcttga ttccgctgag tcaagaaacc tttgcgagtt 180

tgaaggaaaa gctttactag ttgtgaatgt tgcaagtaga tgtggttaca cttatcaata 240

tgctggcctt caaaagttaa atgaaagtta taaagatgaa gattttctag taattgggat 300

cccatctaga gattttcttc aagaatactc tgatgaaagc gatgttgcag aattttgttc 360

tacagaatac ggtgttgaat ttcttatgtt ctcaactgct aaagtcaaag gaaaaaaagc 420

acaccatttt tataaaaaac ttattgcaga atcagggttt actccctcat ggaactttaa 480

taaataactta atctcaaaag agggcaaggt tgtatccaca tatggatcaa aggtaaagcc 540
  
```



tgattcaaaa gagcttatat cagctataga aggcttgctg taaaattatt acttagaaaac 600  
 taatacagtt ttaggcttgt ttgctgcaaa tattccatta tctacaactc caggaatatt 660  
 attaatacaa gcttccattt cagtggggtt tgaatatatcc atattagaga tatctaaaaat 720  
 gtgattacct tggctctgtta taaatccagt tctatatgtg ggtattccac cgatcgagat 780  
 tatttttctt gcaacaaggc tcctactttc aggtatcacc tctataggca gtggaaaaagc 840  
 tcccaaaaaga ttaaccatct ttgactgac aactatacat ataaactcgt tagaggcaga 900  
 agcaactatc ttttctctag tatgtgcgcc accaccacct ttaataagac aattttcagg 960  
 agacacctca tctgcaccat ctatgtaata agctatatca actacatcat taaggctaaa 1020  
 gacctctatc ccattttcat ttaataattt tgatgaagca tctgaactag aaacagctcc 1080  
 agcaaatattg tgcctatgct ccttttagttc ttctataaaa aaattaactg ttgagccggt 1140  
 tccaataacct aaaatcatct caggatgaag attatttttg atatattcta tagctttgttt 1200  
 agcaacattt atctttgagc cactcataga gttataatac aagaaaaatat aggtagttaa 1260  
 ttatttttag actaaaaatt aaaaaaacag gttcttttta gaattcccag aagtacctaa 1320  
 agcttattcc taatgcatct gtttatgacg tagcattaaa atcacctata acatttgctc 1380  
 taaatatattc ttcaaagctg gggaataaaag ttttcctaaa aagagaggat ctgcaaaccta 1440  
 tatttttctt taaaaacaga ggagcgtata acaagattgt aaatttatcc gatgccgaaa 1500  
 agaagagggg ggttattgct gcatcagcag gaaatcatgc tcaaggggta gccagtgcac 1560  
 gtaagaaaatt aaaaattaat tgcttgatag ttatgccaat aacaactcca gaaataaaaa 1620

09847513\_0000001

taaaagatgt aaaaagatgtt ggagccaaaa tactccaaca tggggacaac gtagatgcag 1680  
 cattaaaaga ggcactgttt attgcaaaga aaaaaaaaaatt gtcttttgtt catccttttg 1740  
 acgaccctct aacaattgct ggccaaggga ctataggaca agaaattctt gaagataaaaa 1800  
 ataattttga tgttgtcttt gtcccggttg gaggaggagg tattctagct ggtgtatctg 1860  
 cctggatagc acagaataat aagaaaaataa aaattgttgg tgttgagggtt gaggattccg 1920  
 cttgtcttgc tgaggccgta aaagctaata aaagagtatt tttaaaaagaa gtgggcctct 1980  
 ttgctgatgg ggtggcagta tcaagggttg gaaaaaataa ttttgatgtt attaaagagt 2040  
 gcgtagatga agtcattaca gttagcgttg atgaggctg caccgctga aaagatatct 2100  
 ttgaagatac aagggttcta tcagaacctg ctggggcatt agcacttgca gggttaaaaag 2160  
 cctacgcaag gaaagttaaa aataaaaaaac ttattgctat aagttctggc gctaattgtaa 2220  
 atttccaaag acttaatttt attgttgagc gatcagagat tggtgaaaaat agagaaaaaaa 2280  
 tattaagtat caaaatccca gagataacctg gaagttttct taagctttca aggatgtttg 2340  
 gcagctctca agttacagag tttaactaca ggaaatctag cttaagcgat gcatatgttt 2400  
 tagttggtgt tagaactaaa actgaaaaaat catttgaaat cttaaaagtcc aaattaaaaa 2460  
 aagcagggtt caccttttagc gactttactc gaaatgaaat atccaatgat catctgaggc 2520  
 atatggttgg tggcagaaat agtgactcag gctctcataa caatgaaaga atatttaggg 2580  
 gagagtttcc tgagaagccg ggcgcgtgt taaattttct agagaaattt ggaaataaat 2640

09847513\_000001

ggmatatttc cttattttcat tacaggaacc taggttcagc ttttggaagc atattaattg 2700  
 gcacgagag taaggataaa gacaagctaa taaatcattt aaataagtca ggnactattt 2760  
 ttacagaaga aacctctaac aaggcataca aagattttttt aaaatgaaag gttaataactt 2820  
 taatctaaat ttaattgaaa aaagctcatc gctagggttt tcccacggct ctttgaacaa 2880  
 ctccgattga gatctatcat cctcctcgtc gtaaatcttc ccacctttag aatagaccaa 2940  
 aaatagatat gacaaaggag cgagctcata tttatatcta atttggaacg aagctacgcc 3000  
 agtattaaat tcattaaacta tattatttcc tttataaagg tatccattag catctgaaaa 3060  
 aatactaata ggatttttttg ctttcaaagc acaaaattga ctcttaagtc tgatctcatg 3120  
 tttattattt ttaaaccaat ttagatcaaa agataaggta tcttgtcttg aatcatatga 3180  
 ggcaagatta ttattatcct gccatatcag ccattcattt tcctttctta ttctatatg 3240  
 cgcattgatt cttaagttat catttggaac tattgaaacct gctatcttgt aaaattttct 3300  
 accataccca ttcgaatccc accgattatc ttctctccc ttaaagaagc taactctcca 3360  
 gtcatatgac cagaatgaat agttcttttg ctcaaagtct gctgtaatac ctattcttct 3420  
 ttttgatttg ataaaggggt aggccttcatt ttttcttttg atagttgtat ttttcccaga 3480  
 agatctaaag ttaaaatcta attgaaattt agagttgtcc ttaaaaactaa aagaattttt 3540  
 ttgatcgtg cctattggat ttgaattacc agctgtgtca gcatcataat ttagatcaat 3600  
 tccataatct atttgtttta atatcgagct attatcaaat tcattttattt ttcgatttcc 3660  
 tccaatacca gcatgaatcc agtctcttct ttgcagataa ccaaagtcac ttaactcgaa 3720

09847513, 0000001

gtcgtcttca aaataaaagaa ggcttccact tatgttagat agttattttg gaagatatgt 3780  
 aaactgagtc ctatacccaa gccatttttt accatctttt tctgaggcta acagatcttga 3840  
 atatgtaatt aattttttttg aacgaatatatt gatgtaatca ataacattga ctgtttgatga 3900  
 ttgcgcctgc atttcattct caacattcgt taccatgaag ccaagcggttt tatttccaag 3960  
 ctttgttcga gatcgaagag cataatagtc tcttccaact gaaaaggctt catcagcctc 4020  
 acttgctaca aatactccaa attcattatt attacttttt tgagtaagtc ttaatgcaaa 4080  
 atcaatatca gaatagtttt ttttagctgc ctcgcaacct tcttcattac tttcttctga 4140  
 gcaattatag ctgggggcag ctccaatcct ccgcgtattt ataaccgagt atctatcata 4200  
 attactaata tcaaatagtg attgggtttc attgaaaaat gctctttttt ctgagtaaaa 4260  
 agtttcttga gcagaaaaagt taataaccac atcatcactc tcagcttgtc cgaaatcttg 4320  
 attaatagct aaattttatt gacgaccttt tccagtgcta taaaagattt cagccccaat 4380  
 atctgacctt tcttggtttg taactgaatt tttatttgaa gatatatatg gaaaaaaagt 4440  
 aagctttgat tttgtatagt tttgtatttc taagctatct aactcttgaa agtagtcatt 4500  
 tctactagcc attgttccgg cactgctaac ccataactca ttttcggcac tataacgtaa 4560  
 tgcggtgtaa ttaatttttc ttatatcacc atcaggctgt ttcattaaag ttacatccca 4620  
 aggaataaaa aactcagaga cccaataccc atcaaatatt tgtgtttttg caatccaatc 4680  
 tccatcccg tctgtcttaa agtctcctgc ttgcgttttt atggcatcga aaagcgagtt 4740

09847513, 0000001

cccaagattt atagcaagaa tgaagccttt gttaccatca ccatcaaagt ctatatattat 4800  
 agagttttta tcgctaagtg aatttatttg atctctaagc gtcttcctgg agaacaataga 4860  
 atcattactt tgaaaaattct taaatccaac atatatacca tccttatttg agaaaaattaa 4920  
 agccgttgta agaagttcat ttttcttaag agtaaaaagga gatgtttcat aaaaatctgt 4980  
 aattcaaat gcattattcc actcaggctc atcaagagag ccatcaataa caattgagtt 5040  
 tgaccaaacc agcacagagg taagtaaaag tgataatgag gctaaagtta ttttcataga 5100  
 tagatttaat tcagagtatt ttaatctatg aagagtagga aatcatccga tcattctcag 5160  
 aaaaaacata atcatattta aggtcatgtt tttctccaaa atgcatatta caaatctggt 5220  
 attgataaca cagtccaacc aataaaggtc ttgataacctc ctggtttata gagccaataa 5280  
 ttctatcgaa atatccagag ccatagccaa gcctgtatcc atttaaatca actcctgtca 5340  
 taggaataaa cattaaatca atttcattta tgttgacata atcttcactt ttaacctctt 5400  
 tgatcccaaa ttgatttata aagaagtttag gctgctcatc cagcaaaatta aaagccatca 5460  
 tttcgtcatc aattactttt ggaatataaa tattcttttt aagtttagta aaagcttgaa 5520  
 ttaacaaaatt tgtattgact tcatttcgaa aagggaaaata taaggcaatg ttttgcattt 5580  
 catgagtatt gattttttct aatacatttt cttggattaa aaaacttata ttgctcttag 5640  
 ataaatcaga aatagactgc ccctggtcaa agagtgattt tcgtatttta ttttccacca 5700  
 taaattgtgc cgaggctaac aaaccaaacc accgccatgt gctataacctg aaccgatggg 5760  
 tcagggtggtg aatttcataa caaatcaggc ttccctgtaa aggtactgcg caacaatgct 5820

09847513 .000001

agaggaatat cacctatttta attgtatcgg ttcaaatTTTA ttaacacatt agcgtatgaa 5880  
 ccagaatgca gttaattata aaatatatat gtattaagta aaagttaatt tttagagagc 5940  
 agactctatt ttttgtatta gctttttcgat atctttatta tttaaagcag agtcagtatt 6000  
 ctcggttggc ttggataaaa gcttactagc catagttaga ccagctatca ccaaggcatt 6060  
 attcttatca ctgatgccat caagctcatc atttaataat tttgcagctc taatgagttg 6120  
 gtctttctct tctggcggac aggctaacgt cagatctcgg ccaaaaattc ttaaagatag 6180  
 cgtttccatt tctgacatta ttttttgagt ttagttattt ctttatttag tatgagctta 6240  
 tcttttttcc agtctctttc cctttgcaca taagaatcta ttattccttt ttgctcatca 6300  
 aatctttttta tcaaaagatc taccttatcc tctagttcaa ataatgatga ttttttttct 6360  
 tctgacattg tataagtta attgtactag aatgaattgg aaagtttctt taataaaaga 6420  
 gtaaaataag gcatggaaaa aataattttt aaaaacagaa gagattcact tataaaacac 6480  
 ctacctaaaga attcagcctt aattgtacct ggtgcagatt tgcaatatag aaatgctgac 6540  
 tcactcttata atttaagaca ggaaagtagc ttctactatc tgtctggctt ttgtgagccc 6600  
 tcttctctaa tggtttttagt taataatgga aaaagcattg attcaataat ttttgttctc 6660  
 gaaaaagata aacttaaaga aatctgggat ggttatcgag ctggccctga gggcgcaata 6720  
 aatgattttc tttttgatca agcttttgaa aataataaat cagatgcttt aatgcctgaa 6780  
 atccttcaag ggctagaaaa agtttttttat tcaataggga agaaaaatgg ctttgatcag 6840

09847513,0000001

aaagtaattg actggacatg cgcagcaaat tctaaagata ggcacagcaa atcaattgat 6900  
 attattgatg gctcttcgat ggtaggaaat ttaaggctta tcaaagataa gcatgaaatt 6960  
 gatattatga agagagcttg tgaatttca gctgaatcat atattgaggt catgaaatct 7020  
 ataaagcctg gggacaaatga gcaggaaata gaggcgctat tttatatga attcgccaaa 7080  
 agggaggaa ggtttccagc ttatacacct atagttgctg gaggtgaggg tgcttgtgta 7140  
 ttgcattata ttgaaaaatga taaagagtta gcttcatcag atttaatttt ggtagacgca 7200  
 ggatgtgaat acaaaaatgta tgcattctgat atcacaagaa cctcccagt aagtggaaa 7260  
 ttttcagatg aacagctaca aatttataat attgtccaca aagccaatct tgctgcaatc 7320  
 gatgctgtaa aaactggtaa tagcataatg gagcccaaa tggtttcaga aaaagtaatt 7380  
 actgaaggtc ttgtagagtt gggatttcta tctggcgatg ttaatcagct tcataaaaa 7440  
 ggtgcattca aggactttta tatgcataag gtgggacatt ggcttggact tgatgttcat 7500  
 gatgttggtg actacatgga gggagatgag tttatgaagt ttaagccagg gatgataacc 7560  
 acaatcgagc caggcatcta tatcagtagc gcaatggatg tagatgacaa atggaaaaggc 7620  
 atcggcataa gaatagagga cgacatcctt gtaacagatt caggcaatat taatctaaca 7680  
 gagaagggtc catctaattcc tcaagaaata gaatcattga tggcttagac tatggagggtt 7740  
 ccaattgtta tttctggcgg agggataata ggtaattaca tttctcttag gcttgaaaaa 7800  
 aataatatca aaaccgttat tgtcgaaaaa gctagtagtt tcaaagccct agataagggt 7860  
 ataagaacag tcaactctcaa tgagcattct atgcaaatgc taaaaaatat tggttattgc 7920

09/847,513, 000000.1

ccatcaattg ctcaaatcaa cagcatcgac gtattagatg gtgagggtac aggcaaaatt 7980  
 caattttctag caaaggacgt aggcagcgaa aacctttcat atgtaaccta tttcaatgaa 8040  
 ttacaaaaaac taattttctga tccatgtaaa gaaagaacct tatttaataa tgagattgat 8100  
 tcagttcaga atcttaatac agaattctgat ccagagatca tgcttaaaga tggcatgacc 8160  
 ataaaaacga atctaattgc tggatgtgat ggaagaaatt caaatattgc aaaaattgct 8220  
 tcacttacaa gcagcttcga tgactactta caaacagctt taacttttgt cgttgatatt 8280  
 gataatgatt cacatggcaa agctcaccaa gttttttctg aaaaaggaat attgcactt 8340  
 atgccactcc cagaaggcaa gggtgagatg aataaatgca cagtgggttg gtcaataaaa 8400  
 aatcaagttt tgggagatga gcctgtatct gagtttgtaa aaaataacat ttcttttttt 8460  
 gaatcaaaag ttaatgttag tctcagggtt aagtcagaaa ttttaagttt taaattatcg 8520  
 aaccatcatt ttgaaaaacta tattagcgga cctattgttc ttcttggtga tgctgctcac 8580  
 tcaattcacc ccttagcagg tcaaggtatt aatctaggat ttgcagatgc agatactttt 8640  
 tgtgaagagg taattagttc ttataaaaaa gggattgcct ttaatgagaa atcagtttta 8700  
 aaaagatatg agattagaag aaaaagtatg aactttttta tgttgaagtc tatggacttt 8760  
 tttgtggatt tatttggttc agaaaattta tatctaaggc tgatcagaaa tttaggtatt 8820  
 tcttcgttaa ataaatcaaa gtttggttaa gcattcttta taagacatgc ctctggaatg 8880  
 aataagtttt aaatttgtat taaacttttt gaccttttagc tctaagttct ttaagaactt 8940

09/847,513,0000001



cactaatgcc ttttttatca atgattctca taccttttgc agatacttta agatttacga 9000  
 acctgttctc agattcaacc caaaatttgt gtgtgtgaag attaggaaaa aactttcttt 9060  
 tagtcctatt tttagcgtga gaaacattgt ttctgactg tggatatcta cctgttacttt 9120  
 gacatatatt actcattgaa acgcgatttt atagaacact gaggaactta gcaatactat 9180  
 tgtgaaacaa atttatttat tacggcatgc acaatctgac tgggagagct ctaatcagaa 9240  
 agatttttgat agaccattag caagaaaagg cattgaagaa gcaataaaaa tatcatgtta 9300  
 ctgcaaatct cattcaattt tagtagataa aatatcttgt agcactgcag aaagaactaa 9360  
 gcagactttt gatatatgca gtgatgggct taattatcca atagctgaag cagtctatac 9420  
 tgatgagctt tacttttctg gccctggtga gatcgttaag cttatccaaa gtttaagtga 9480  
 attcatttcc tctgtttttaa taataggcca caatccatca atgcaaatgt acatagatgc 9540  
 tatttcagaa aatcctcata ttacgtattc aacatgcggg ctggcagaaa ttctcgttga 9600  
 aagttcatgg aaagacttat ctttaaaaaa atgtaagta aaatctttta ttcaaccagg 9660  
 agagctttta aaattgaaaa acataaaaaa taagatcatt aaccactaa tgggatccaa 9720  
 gataccctta cctcaatatg aaacaaaagg ctcggcagga ttggatttaa gggcatgcct 9780  
 agatagtaat ctcagccttc aagcaggaac atctcagttg atacctattg gttttgcaat 9840  
 gtacttagaa gatcccggtc ttgcagcaat ggttatacct agatcagggt taggttctaa 9900  
 gcatggaatc gttcttggtg atctggttgg gttgattgat tcagactatc aaggagagct 9960  
 aatgggtcct gcctggaata gatcagatac agattttgag attaatcctg gagacaggat 10020

09/847,513 DeLong et al., 2001

tgcacaaatg attatagttc cagtgattca agcagatttt gaaattgtag acgagttcaa 10080  
 tgagactcag aggggagaaa agggttttgg aagttcaggt ataaattgat aaatttactt 10140  
 tttcttgcca aatcttttctt caaattttctg aactcttccg ccagtatcaa taatttttttg 10200  
 cttaccggta taaaaaggat gagaagcaga ggatatatca agagggtagt atgggtatgt 10260  
 ttttccatct tcccattctt tcgtttgagt cgtatctaata gttgaacgaa tgagaaaagaa 10320  
 cttatcagca ctacgctcat ggaataaaac ttcacgggtat tcacgatgta tatctttttt 10380  
 cataataaaa tcaaaatttg gatgagaact ataacaaaaa acaacttctt ttcaatcaaa 10440  
 acatgaaaat tttttactat gacatagctg tttcattgcc tctaaggcaa tgctttactt 10500  
 atagtttctga gcttaaaaatt acaaaaggaa cccgcgtgtc agttcctttt ggtaaaaagga 10560  
 agattgtagg agtggttata aaaaatatctt aaaagccaga tttcctaata aaagcgggag 10620  
 ctataaaaaa aattatcgct gtacttgatg aatatccttt gttcgacaag ccaattttttg 10680  
 attctatatt gtggtcttctt gattactatc atcatcctat tggtgaggtt tttaatatcct 10740  
 ttataccccc cgaattaaga aaaattaata ataaaaaaat tgaagcttta agagaatttt 10800  
 ctgaatatct agtaaatgag gatgataaga aattcgattt aaccaaggat caagaaaaaag 10860  
 cagtcaaggc gcttttctaaa tctaaaggat tttcacccac tttattatat ggagttacag 10920  
 ggtctggaaa aacagaagtt tacttaagag ttgccgaaac ttttattaaa aataataagt 10980  
 cagtattagt tttagttcca gaaataaatt taacccccca attgctctcg cgatttgaga 11040

09/847,513 . 0000001

atagatttaa tgggtgaaatt ggcataatc attctaagca aacagcagct aagagattaa 11100  
 agacttggct aaaagctaaa ttgtgttcta taaaaataat agtaggaact cgatcttcgg 11160  
 ctttagtgcc ttagagataac attggtttta taattatcga tgaagagcat gaccaatcat 11220  
 ttaggcagtc agaaggggtt aaattctctg ctagagactt aagtataaaa agggcacagc 11280  
 ttgcagatat tccaattatt ttgggatcag caacccttc gctgcaaact ttaaaacttg 11340  
 taaaagaaaa taaatttata agagttgata ttcctaactg agttgatgga aacaagcctc 11400  
 ctaaattaat agccttagat atcaataaca gccctttaat aggcggagtt gctaaagaga 11460  
 caattgaagc aatgcaatca accatagaca gaggagaaca ggttctagtt ttatttaata 11520  
 gacgaggatt cgctccactc tatcaatgta gtagtgtgg gtgggtagca gattgtaaat 11580  
 cttgtgatac aaatttagtc tccaccagg caagaaatag attaatgtgt cataggtgtg 11640  
 aatctgccta ctctgttaat ttgtcttgtc cggcatgcaa gtctaataatgt ttaatatatgt 11700  
 atggagctgg aacagagaga gttgaagaag ttcttaaaag cagctttgta aagactccaa 11760  
 taattagagt tgatcatgac tcaacaaaaa aagtgggagc tatggaggct atagttaaaa 11820  
 aaattcattc ctgagacgca gcaatttttag ttggaactca aatgcttgca aaaggacatg 11880  
 attttcctaa agtcacctta agcgttatth taaatgctga taatggcctt ataagcccag 11940  
 aaattaatgc attagagaaa atatctcaat tgcttattca ggtctctgga agagcaggaa 12000  
 gaaataataa tcttgcaaaa gttattattc aaacaagata tcttgatgat ataaatctta 12060  
 ataaaaataa gacaggagat tatatgaaat ttgcttctca atgcctaagt accaatgagc 12120

09847513 DeLong et al.

aaatgaactt acctccattt actacttttt gtctgcttag gtgctcatca ccaactcaaa 12180  
 agagtaatgt agatttccta gagaaagctg ttttaatttt atccaatagg actgatataa 12240  
 atgttattgg tcccttgctt tcattagttt cgaagtcgaa aggaaattat aggaccaag 12300  
 tctatatcca tgcaccaaaag aagacttttt taaataaggt attaaagttt ttgacaacag 12360  
 agtttgaaaa atggccggaa tctaataagg ttaagtgggc ttctgacatt gatccaatag 12420  
 acttaagcta aatattaatc ttaattaatt gtccctgggt tattgggtta ttgtttagtt 12480  
 tattctctgt attaatttct tctacagtca ccccaaatct tctcgctatt tctgataaga 12540  
 catccccctt ttgtattttg taagtcacaa agcctggatc aatactcata aaggtatttg 12600  
 ctttaggttt gtcccttaaaa tagttatgta ttccctaggaa aacagatctt gcaatcattc 12660  
 ttctccctgg cttgcccctt aatcttttgt cgtcttcagg gttaggttata aaccctgact 12720  
 caaccaatac agaaggtata tcaatagact taagaactct gaaatcagcg tactcaacat 12780  
 tcttttttat aattttttgt aaagggtctc ttttaagttg atccaatata ttagttccta 12840  
 aaattttact ttcagaaaatt tttttcttat atatttctgg ataggtttgt cttgccgcat 12900  
 cctcatcaa atcaactggc tttagatttt ttatatcagc ctgtattctt tcccttttgt 12960  
 ttttagataa gtctcctagca acagtacttg aagcttcac tgaccatata aaaacagaag 13020  
 caccctttac ggaagataat ctaaaacccat cagcatggat tgaaacaaaa atatctgtc 13080  
 catatttctt agcatcctgg tatctattat ttaaatctaa tgtctcatct ccattctctaa 13140

09847513,000001

tcatcaccgg tctgtaccga taagtatctc ttaaggttct ttctaactcc ttgcgaataa 13200  
 gtaacgtcac atctttttct aaaatattat ttggggccaac cgcaccagggt tatttaccac 13260  
 cgtgaccgcg atctatagca acaacaatat ctctaatact ttattagggt tttttatttc 13320  
 ttttcacctt aagctctaatt tttatatattt ctgtatttat agtctgagtt ggggttttgcc 13380  
 aatggactga ttcatataaa tcaacaacga ttctgggtaag acttcccatcc tgagatgctc 13440  
 ttactttttt gattggatag ttgtatgga catttatctc ggtcggaagg ctcgattgat 13500  
 taattttccat gacaattcta gagggatttt taaagaataa gaccttacca atgaaaacttt 13560  
 atctagacta aagttaatac taatttcatt gttaccatgt tcttgtattt catcgaagaa 13620  
 aacatcattt ccgctgataa aaaaagatag aaaaccaagg attgccaatc tattcataag 13680  
 tttttaacca ttcttttaaag ccattgtcac cagaaattaa agaaacctct ctctcctcag 13740  
 gaagatggct gaaaattatt tttagatcaa aacttctttc atgttgaagc ctttcaggcc 13800  
 actcaattag cactacttct ttagagtta tttttctact tagatcaaat atatcaatat 13860  
 cttctgcttc gtaggttcta taaagatcaa tatgcaaaaa cattaaatta ttaaaatcat 13920  
 attcctcgca gagagtgtaa gttggacttt ttacaagatc ctcccatcca caattttttta 13980  
 taatagatct agatataaaag gttttgcctg ctccctaagtc tccctcaaga tgtattttcaa 14040  
 tttcttgcca agaagattta agtatctcca tagctatttt tgaccctagt tgggtttgtag 14100  
 cctcgtcatt tataagagta agttttttca tcgatttatt aattgtctta ttataggat 14160  
 taaactagag gcatttaagc ctatttcacc aatatctacc ttaaaattta aacctgcctc 14220

09847513\_0000001

cgaatgaaca gcaacagcaa taatgctgcg attttctata tcaagacctt gggcaagaag 14280  
agcagtcaaa acaccagcca gaacgtcacc agttcctcct gaagcaagtt caggtccgcc 14340  
gcacgcgcac ataaaagact tgttatttgt tgagtcaaa accaccgttc ccattccttt 14400  
taatatgaca atagaagccc cataagtgtc tgcaatttgc ttggcagcag aaatcctatc 14460  
tctttgaact tcttcaatag atatgttttag taatatagcg gcctctcctg gatggggcgt 14520  
catcaagatt gttttattgg attttttgat taaagatttt gatgatgcca cgatgtgtaa 14580  
agcacctgca tctaataataa ttttttgaatt atttgcagac tttaatatatt tacccaatat 14640  
agtttttgca aaagcagtat tagcaattcc tggtcctac aatataaacc tatgatcttt 14700  
aattttttaa tctatatcga ccccttcaac catatcaaca ccaattgcca ttacctcagg 14760  
gtttctctaat aaagatgggc ttacattaga cgtgtctgtg acgagagtta ctagtccaga 14820  
gccacaaaat aaacttgcct cagatgcaag tattccggcg ccacctatgc ctggggagcc 14880  
agcgcatatc aaaacttttc caaaattacc ttatggcta tcctttgctc tattgggcag 14940  
aagatttttt aattcttgaa aggtaaaaact ttgcaatata gacatgaaat tacttaagct 15000  
aaattatgtt gaatgctaaa aattcattct tcagcgaaat aaccaatcct atcggaatgg 15060  
aatgatccca taaaaatttt ttctgtcatga ggcactgcaa tagtgggtag gccaaaaact 15120  
gttctagaaa atgaaaaactt atcttctctc ttaagtgttt ttgtattaag tttatgaatg 15180  
gagaagggtg gagaacaatt tgttttttct gcacaattac caaaatcatt agcctgtaaa 15240

05847513\_099880.1

tctaagaag ttaaccaaac cgaaccgtct tttaaaaaaga tattatctgg actctgaata 15300  
aaataactac ctgttttagt attttgatta atatcgtaaa cagataaaatt gtcaccttga 15360  
ttatagttaa cataaagaag tccagagccc tcatctagca aaattccatt aggcccactt 15420  
ccatctgttc catcaacttt attaaaaatta ttgtcactcc aaagtactac atgccctgaa 15480  
atacttttaa ataatgaatt cattaaccat tctttcatgc tgatatctct tttatacata 15540  
tgagatgcat aaaaacttcc atctttttta agggcaacat cattgaaaata atattggtca 15600  
gggaccctaa tacagccccc ccataatcata tcccatgaag attcattttt tataatttca 15660  
aacatttcaa tcgactcaa tggcgaatga ttaattacag ctagttgata gaagcccttca 15720  
tcattttcaa caagatctat tccgtgggga ttgaatatgt ctagctcgcc tcttatgcat 15780  
gaagcgtcac ccaagagct ttctccaaaa gtaatttttg gtaccttttt ttcataggta 15840  
tttaaatcca ttaaggcaaa gtaaccaggg gtatgtttctg catatggacc gatccctcca 15900  
aattcagaaa taaagaagaa tttattatca ggcgtaatca caatatcttc tgggtttgag 15960  
aaattacaaa taaccttgat acggtcatcg gattcacact tactaatatc catttggggg 16020  
cctatatagt cagtagatac gacagttaca gaaataaata aataaagaat agaaaccggg 16080  
acggtaatct tataatagtg ctttataaaa atttctaaaa tctttgaagc atgatttggg 16140  
agtgcgatca gccaaacccc ttttaaaaaa gatagagccc ccataacaac aaaaattact 16200  
gaccatagcc tgtcagacca ttccgggtga agaagaccag ttaaaaaaaag aatcatccca 16260  
aagaataaag ctaaatagcc tgagattttg actttcgatc caacaaattt cgaaaaagaga 16320

09847513 1000001

tatgtataaa gaggcctttat aagaactagt aaaccacagg ccaagaaaaa gaatgctaag 16380  
 tagtaattca taagttagtt ttatatataa tgctccttaa taatactaac aagtttctaag 16440  
 ggcttgcca atggaacatg gtgagcagct ccaggaaccc cctcaaaaagt cataatgtca 16500  
 ccatatgtat ttttaatat gtccaagata ctcccgagg ttaataagct gtcttcaccg 16560  
 tggtatgaaca aagcagggca gccaaatgaa aatgtgtaac cgaataaacct ttcaagactg 16620  
 ctaaacatga catcatcaaa ttcccatctc caccagcct caatatTTTT tactgagtgc 16680  
 tcagcaatgt atcttaagta ccaatcattc gtacaatctt gcttaggcac taacctaaac 16740  
 cttttaataa tatctgtctt gtcttgatag tgcttgatca ttctgagagg agaagagtgt 16800  
 tgattcgggt cataatccgg tggcttata aatgtatcaa taataataat attatttatt 16860  
 agatcccttc tttcagatgc aacgtaacca gcaacatgtc cgccgaggga gtgtccaaca 16920  
 ataaaaatat ttgaaatat ttttttatcc ttttcctttt caattacaga gacaatacat 16980  
 tctccaaaat ctttaatgcc atatgaatct ctaaaagaag agtcacccat gccagggaaga 17040  
 tctattgcaa ctatatattgc gcagtctcta aagtggggcg caataggatc ccaccatttt 17100  
 ttatgagcac ctgttccgtg aataagaatt attaaatctt tgctttcatc ttgggagttc 17160  
 cagctagaat aggatatatc cccatgagga ttcttgataa tctctgagct aggcttgtcc 17220  
 tcaatggcat ctttgaacca ctgtggggca tgaataatgt cttgatttag attgttagtt 17280  
 atttccataa acagtattct aagctataaa aaataaaaat atgaataaac ttaatttaac 17340

09847513.000000



gccagcagca actgttttttag tcctaaagga ttctcctgat gggatggaag ttttgatggt 17400  
aaaaagatca agtaggcctc ccttcggaga ccttttttgtt ttccccgggcg gcaagattga 17460  
cgaaggtgat ttcaataata agatagaaga tttttgtgag ggcgtgactg ataaagagggc 17520  
ctccataaat cttggattag attctggagg tctagcatat tgggttgcatt gtattagaga 17580  
atgctttttag gaggttgga ttttacttgc taaaaaaaaag agtgggggaag atcttgatctt 17640  
agatggagtc gataaacata aatatcaaaa atatagagag atgttgttaa ataataaaaat 17700  
tgattttatat aaaatctgtt tagaagaaaa tttaatttcta atgcctcaac aaatagcccc 17760  
tttctcgcatt tggataaccc ctgaaaataga aactaggaga tttgatatac gtttttttttat 17820  
tgccccacctc ccaaagcattc agaccggaga acatgatggt agtgagctca tagacagtgt 17880  
ttggatttca caaaaagaag cgctcaaaaa atctcgttcg ggtgagatgc ctatgattat 17940  
gcctacaata aaaaatttgg aacaatgtgc acaatttgat tcgggctcta agctttttaga 18000  
aaatcagagg aatctctcaa atgaggatat ccaccaatc ctgccaaagt tttttaaaga 18060  
agatggtgag tggaggggtc tattgcctgg agataaaagg tatgaggatc attaaataat 18120  
atggacttaa ttactaaaat aacagctccc aacctgggtg ttttactgg ggggtgggact 18180  
aatacttatt tgattggcaa agatgatata acccttgtcg accctgggtcc aaatatatct 18240  
gagcatctag atgaaattat caaagcagg gatgggaaaa taaaaagaat ctttgttact 18300  
cataccata cagatcattc ccagccgca ttgcctttat caaaaaactct taatgttcca 18360  
atgtacggaa ggctagtaga tgggtgaatcc tcatgggagg atgaaacatt tatcccagat 18420

09847513.000001

attatttttaa atgataaaga tattatttgag acagacgaat atacgttaga agtaatacat 18480  
 actcctgggc acgcatctaa tcattttatgc tttttaataa aagatacgaa atgcctttcta 18540  
 acaggcgatc acattatgga cgggtctacg gttgttattg ggccaccaga tggcaatatg 18600  
 acaagctata tcaatttcatt agaaaagtta ctgatttttg atattgattg ctttgcgcct 18660  
 gggcatggaa attatatattca tgagcctgag aaaaccattc aatcaattat taggcacaga 18720  
 ctaacaagag aaagaaaaagc tcttagaaaag ctaggagagg caggaatctc atcattagat 18780  
 aaacttacta agcttgtttta tgatgatgtg tcagagatgc tccatcctat agctaaatat 18840  
 agtctagaag cacattttatt aaagcttata gatgaaaaga aagttaaatt agataaaagat 18900  
 ctattcgaaa taattttaatc cttttttattt ttatgtaaga ctttctcctc aatagcttct 18960  
 atatcaatat catcaattga gtcttcgtta ttatcaggta tctttttttac atcttttttca 19020  
 attttaaggt cgataggaga ggctccaaga tcaaaagtta gctcccttac atttttttgaa 19080  
 atagtatcct caacacagtc atcttgatcg tatgcctcac ggggtctctcc ttatcatta 19140  
 atgggaaaagg gtcttttgggg agggcccat tgcatgcatt taatagtagc aactgggtgag 19200  
 tagtaatcgt cactagaata taacttatca agttcttttg gtgatattga gcaaccaata 19260  
 attccaaaata gaataaaatgg cgctaagagt cttttcataa actttgtttg ttttctataa 19320  
 gagactcaac aaccgaagga tcggcaaggg ttgtcgtgtc tcctaaaatta gatagatcat 19380  
 tctcagcaat ctttcttaaga attctgcgca taattttttcc cgatcttgtt ttaggcaagc 19440

09847513.000001



caagatcttc ataccaaaaca tctttttttat catcccaatt aatttcgcca ccagttctttt 20580  
ttataaccag tgtatttttta acatctggac agcccagaag agcctcatct acattagatt 20640  
taagtggcac ttttttgccg ccccttaaac cttcatcagc agttataaca attttacaat 20700  
cagcatcaag aattctatct ttgagtgatt ctggagaaaa gccaccaaaag acaacagagt 20760  
gcacggcacc tattcgcgtg caagcaagca ttgcgaacgc agtctcaatg atcataggca 20820  
tataaataca aactctcgag cccttttgaa caccagggtc ttttaaaaca ttagcgaact 20880  
tacatacttc gtcatggagc tcttttgtagg tcaattctttt agaatcagca gggtcatctc 20940  
cttcccatat taatgctatt ttgtttggat catttttctaa atgcctatcg atgcagttta 21000  
agctaataatt tgtttttacca ccctcaaacc acttcgcatt attaaattga ttattgaatg 21060  
ttgttttgaa gtcttccatc cagcttatgt tttcgtttgc tagattttta aaaaatttag 21120  
aaggatcttc tatggattgc ttgtaaagtt ctttgtattc atcaaagtct tttatataag 21180  
gattacttga atgtttttggg ctataaaagct ttcgaggcat tottaaaataa ttgaaggttg 21240  
ggggttaatg aaattctttc ctttgggatt ggacattatt tttgtaatga gcgattcgt 21300  
atcgctatca ttggttttcta aattaataatc tgcagcatta attattaata aaggggcaga 21360  
actatagtct aagaaaaaacc ttgagtatgc atcattcagt ctttccaggt agtcaagagt 21420  
tagatattgt tcgttaatat ttcctctctt agtaatccctg tcttttaaca catcaatagg 21480  
tgcctgaaga tagattacta ggtcgggtgt tggcggctcc aggggttagat ggtcatatac 21540

tttgtcatat agatccattt cctcgtttaga aagagtaacc tcagcgaata atcgatcttt 21600  
 ttctattaaa aaatcagcaa ccctcactgt ttcaaaaagg ctctctttgtt taagatcttg 21660  
 aatttggtgc attctttgaa acaagaagaa aagctgagt gctagagctg attggcttgg 21720  
 gtttttataa aaattcttta agaagggatt ctctgccggt tgttctaaaa aagaatcata 21780  
 attaaatggt tcggctatct tatttgctaa agtagttttc ccaacacctt tcggtccttc 21840  
 aattgcgata tatttttgaa gtggtacttc tttaatagct gggttcattt ataacgtata 21900  
 gtttttatcc tagattatct caactcgtta tagaagccaa atatttttgc agattcttca 21960  
 tcactttttt cattcatctc agcaataatg gtctctctca atgagtcctc atttttttgg 22020  
 aaatcagtcc accacttccc aataccactc ttgtctccag aagctctctc tcttacaagg 22080  
 agaaaatcat atgcagctct aaaccttgga tgtcgaagag ttttatatgg ctggctacca 22140  
 attctgctat gaagttttaa ttgaagtacc caaatatcct taatatagct tgaaaaatttt 22200  
 cttgggtattg ctgtgatttt ttgttgttcg cgaagtacac catccataga tcggaaaaaat 22260  
 ttctcacaat taattttctc attcttagaa cactttttca ataacgaggg ccataacaat 22320  
 gcagccataa gaaagcctgg tgtaattgac tgttggtttt taaccctatc atcagtattt 22380  
 cttagagcat gtgtcataac attgctggca aagtcattcc tacttggtac agaaaggata 22440  
 agataatttat taaatgaaa tgagcagagt ttttcaaaat tcttttcacc cattccattc 22500  
 aaaaatattt tgcagaactc atcaaaccaat ctagcgtttg agatgcctga taaaagatgg 22560  
 cctttatcat agatggcatc cttaaccaga ttatctatct taaaattgag ttgtttacta 22620

09847513 . 0000001

aatcttatag ctcttaagct tctcactgga tcttcttcga atcgtctttg gggatctcca 22680  
atagatacaa taaccttctt gtgtatatgc ttaagcccat cgttatgac tctattttt 22740  
tttghtaacag gacagtagta aagggcattc acagtaaaat ctcttctatg acaatcttgt 22800  
tcaagagtgc cccaattatt atctctaaga atctttcctg tcgaatcagt acaatatatt 22860  
tcaccatctt cctgatcact tcctgatctg aaagttgcta cttcaagcaa ttcactccta 22920  
ttaaaaaacat gaaccaattt aaatcttttg ccaattattc ttgaagcttt aaatgttttc 22980  
ctgatctgct caggcgtggc attagtggt atatcgaaat cttttggctc tagtccccgt 23040  
agtgcacgc gaacgcacc accaaccaga taggcctgaa aattattttt ttgaaggtct 23100  
tgaacgacag atatagcaaa ttacttattc ttattattat ctatcaaatc taattatgaa 23160  
tttttatgat ttaagggcat tagccattta gctgcttttc tttaatttcg tcaagtgttt 23220  
tgcaatgaat gcaatgagtt gcagttggtc tagcttcaag cctctttaatt ccaatttcat 23280  
caccacagga ttcgcaccaa ccataatcat cttgcttaat ttgttcaata gatagacct 23340  
ttttgctgat aagttttctc tctctatctc tggttcttaa ttcaaaagca aactcttctc 23400  
cttgagaagc cctgtctact gggctctgcat aggtttcacc tttagctcta agatgatcaa 23460  
aagtttttttg catttcatcc ttaagatgtt ctttccaaag aagaagaacg gcaacaaaaat 23520  
gtttcttcat tgctgcactc atatatatttt cacccttctt agatttataa ggtgcaattt 23580  
tagatttatt attagctatt gttgcttttg cagacttctt ggctgcaact tttttaactg 23640

09847513\_000001

gagctttttt ggtaacagtt tttttaactg tagctttttt ggcaacagtt tttttaactg 23700  
gagctttttt tactgtcttg gattttttt cgaccatgta agattttata gaatttttagg 23760  
gtggagaaaa taccagatac tgacaaaatt agctagtcat tttttaattt atttaatact 23820  
ttcaagtacc catccgagct aagtcttggc ccaaatgtct caacaacctt ggaagatgca 23880  
tagtttgcaa acttagcaca tgcttcaata ttattccctt gaaggtaggc atgcataaac 23940  
gatccggcaa acatatcacc agctccattg gtatctattg gagttatttc ttctgcttga 24000  
gcatgtttct caaccccttt atctataaca acacttccat cggcgccctt tgtaatagca 24060  
gtcatatagg gcttttcttt ataaaagcta acagcatcat caaggctttc ttaccagaa 24120  
aaagcaacag cttcatcatc attacagaag atcatatcta ttccatatga ctctattaaa 24180  
tcaaattttt ctttaaaacc atgaacaata cctgcatcag aaagagacaa ggcttttcttt 24240  
acgtccctgt ctttaagggt ctctaagact gaaataaacag cattaagtt atcgtcactt 24300  
gttaccatgt agccttcgat ataaaaaatt tttgaatttt ctacaacatc aaaatctata 24360  
tctgatttac caagatacgc actaactcca agcatgctgc tcatagttct cttagcgtca 24420  
ggagtaacta aaattaagca tttcccagtt ggttgatctg tattttcaga gctgacacca 24480  
atatgtttga ctccagccga cctgagacta tcaagatagt ttcttccatc ttcatcatca 24540  
gaaactctgc atacatgatg gcaattcgaa ccataatttg ctgcagcaac aagagaattg 24600  
gttgcacagc caccgcaatc agaaatcgat tcggctccca tttcaataag tttagctaatt 24660  
atagggtgct gtcttctcaga agatgaaaga gtcatagagt cggctacaag gcctacactt 24720

09/847,513 DeLong et al.

gataaaaaatt catggctgac tttatatattga gtatctacta aagcatttcc aagggcgcta 24780  
atatcatatt tcatgtgtta ttccctttttg tattatttgt tttactgttt ctaacgttct 24840  
attaatttct tgatcccttat gcattgctga aatgaagcct gcttcataatt tagagggggc 24900  
aaagtatatt ccactttctaa tacatgaatt taaaaaattt gaaaaatagta catcatcagt 24960  
ttttgcaaca tcattaatat tattaggag ttctttctgaa aaaaagaatc cgaacattcc 25020  
accaatttctg tttatagaaa acggaatacc tgactcaatc attaaggttt tcatcccatc 25080  
caaaagcaca gatgcatttc tttctaactc tttaaacgga ttttctttta tcaacaattg 25140  
caacaaagca gttcctccag ccatagtctag tggattgcc aacaaagtcc ctgcttgata 25200  
gacaggacca gaaggagcta gatagttcat aatttcttct ttgcctccaa aagctccaac 25260  
aggaagaccg ccacctatta ctttccccc aaagcagtttaag tcaggagtaa tattataaat 25320  
ttcttgagct ccgcctagcg aaactctaaa gccgctcatg acctcatcaa atattaaaaat 25380  
agaattattg gctgagggtg tttccctcaa taactttaag aaatcttcat gacctggaac 25440  
aaagcccata ttccctgcta ctgggttcgac tattacggcc gctaagtcac cttttatctc 25500  
attaaatatt tctaaaaact gttcttttatt attgtattcg caactaaatg tatattttgc 25560  
caaattctga ggaacccctg gagagtcagg taagccaaaag gtggcaaccc cagaacccgc 25620  
cttaattaaa agagagtcaa catgaccgtg ataacagcca tcaaatttta taattttatt 25680  
ccttcctgta aaacctctgg ccaatctaatt cgtgtctatg gttgcttcag tacctgaatt 25740

09847513.000001



aaccattctt attttttcaa ttgaaggaat gcatttttta attagcctgg ccacatcaga 25800  
ttaagactt gtgggggcg cataactagt tccaagcgca acttgatttt taattgcacc 25860  
tacaatgtct ggggtgtgaat gaccattat catcggaccc caagacccaa tataatcaat 25920  
atattcatta tgatcagcat cataaaggta tgcgcgggat gctctttcaa aaaatatagg 25980  
attgccattg atatttttga atgcccttac tgggtgaattt acccctccag gcataagagt 26040  
cttggcctct ttgaataagg ctatggattt atcaaatttta ttatgtgtca aagctaattc 26100  
ttcttaattt ttaatttcga tatgatatca acttaattct tattatgtat ttaattgttt 26160  
tgaaaattca tccacagtat tccagttggt gaactcataa gtattttgtaa catctgtagg 26220  
acccttttga atccacataa tcaatctaata catgtgttta tctataaatt tatatttttg 26280  
ataatctatt ttgcctgcaa agactgctaa tttcattggg ttccaaggag acagctctaa 26340  
aaatttttgc atatacgggt ttgttttcagg tgtatttttt tcaggccttc tcgcaactac 26400  
attaactgaa aagaaggcat tttcttttgt ttcaagacaa gcaacatttt tttgaataaa 26460  
ttcataaaagt tctggtttgt gtttgccata cctaattgct gcaccaataa taattttatc 26520  
aaattgatat aaatctaact ctactgcttt tgcaatatgt attatttttg aagattcaga 26580  
cacatctaata ttagagaaaa tttttttaca aatttccaga gtttgcccat cggttgtaga 26640  
gtagattagt agagttgatt tcataaatat attttatatt gttaattaaa aaatttcattg 26700  
gcacgcgcaa aatataaaca aaaatatagt atatttaaaa aatgattaaa tatttattta 26760  
ttatcgcttt attcatcagc aactttgcat atgcaaatat tgatgcagct aggtgcgcag 26820

ggattagtgc agatcaagaa agacttgatt gctatgattt aatttttaaa gctaatgatg 26880  
 agctaccttt agatagtaat attaaaaactc tcatcacacc agctatcaag gctgtgaccc 26940  
 cagctgattc aataaaaaatt gaaaaataaag caacaaaaga aaaggatttc gggctttccaa 27000  
 aaacaaaaat caaaaactct gcaaagaact caataaaaac ctcggtagta aggattaaaa 27060  
 aaacaaaaag tggtaaatta atttttactt tagaaaaatga gcaagaatgg actgctgaaa 27120  
 cttctttatag agcaaggaat atgttttaaac cagaaaaccgc agtcatttta gaagaggccc 27180  
 tggttagtagg tttttatatg attaatataa gtaataaaca gaaaattaga ataaagaggt 27240  
 tgaaatagca atgaccatac agagtatagc aacgaccgag ggtgccatga aaagaataag 27300  
 atctgttctg ggggtcccaag atggctcaag cttcagggtt tatgttaccg gaggtgggtg 27360  
 ctctggggtt caatatggct tcaagtttga taacgatatt gcattcgcag atgatgttat 27420  
 aaattgtgga gatttctcgc ttttaaataga ctctatgtcg taccctctatc tttatggatc 27480  
 aactctggat ttgtgtgaag atctctcagg ggctaaattc gttattaaaa acccaaatgc 27540  
 caaaacaaca tgcgggtgtg gagagtcatt tacagtttag attttgtaat tgaacccaaa 27600  
 agtcctttta cattcttgct tgatgtctga actatcagag cctcattact gactctttta 27660  
 tatcccatcc aggcaaatgc catggactca atagcaaaaa catcatggcc taaatcactt 27720  
 gagaggacga tatcattact agccatctca gaaattcttt ttaccagata ctgtttatgg 27780  
 gcaccaccgc cacaaatgac aatgtcacia ttattatgac cattttttatg gattgaattt 27840

09/847,513,000000.1

attatagatt ttgctgaaaa	ttcaaccagg gtacatagaa	tatcttcagc	tttctttttt	27900
aagaatttct tggataatat	ttttatatata aagagctctt	tgccctgttga	ttttggacat	27960
tttctttttaa aaaaattatt	ttgaagcaat cttcctaact	caatgtgatc	tactttttcct	28020
ttggctgcaa ttgcaccatt	tttatcaaaa ggaatctgaa	gaaaatcact	acaatatgca	28080
tctaaaaattg cattaccagg	gcctacatca gttccccata	tgtcatttct	atttttttaca	28140
aatgagtaat ttgaaatccc	ccctatatat aaaataattc	gtgggtttct	agcttttataa	28200
aataattgat tatggaattc	cggaacaaga ggggcgcctt	cgccacccag	agcaatatgc	28260
atattttctaa aatcacttac	caccagaagg cccgttttct	ttgcaacaat	atttggatca	28320
ccaatttgca tagaaaatgg	aaatctctta ttaatttcat	gcctaattgt	ttgtccagaa	28380
atagcaaac accttataga	tgattttctta attttactaa	atccaatgc	ctcattaata	28440
gattttgaaa ataagaatcc	tatctcttta ttgatagtgc	ctaaatctga	caacgagctt	28500
tcgttatatt caataagttt	ttttactttt aatcttaggg	attttgggaa	tttaattgaa	28560
tggaataact caagatgtat	ttttgtacca atgctaagaa	atgaaatata	aatagcgtca	28620
tgactagttc ctgtcattgc	tccaatatag atcttttttac	tcattactac	ttggagaaaag	28680
tttattgaac tcctccattg	ataatttatt attctttaat	agtgtatcaa	aacttgcctt	28740
taaaattttt ggtactgggt	cagctgatgg taattttact	tttataggat	ctattctttt	28800
cttacctatc ttaaactcat	aatgtaaatg aggccctgtt	gccagtcctg	aactcccaac	28860
gaaaccaata gtatcgccct	gagaaaacttt tttagcctttt	cttataccct	tactaaaattt	28920

09/847,513 DeLong et al. Corrected Sequence Listing (August 4<sup>th</sup>, 2001)

ttctaaatga caatatctag tgaataactc atttgtatga tttataacta tctcattacc 28980  
 acatccattc ctttgaccag aaaaagaac aatcccatcg cctgtagttc tcaccggcga 29040  
 tctctttttt gctgcataat cgactccatt gtgagctctt attgtatgaa gaacaggatg 29100  
 cattctgttt ggattaaaaat gagaactaat atatgcaaaa tctaaagggg ctcttaaaaa 29160  
 ggcctttttg atattatttc cattttcatc aaagtactgt tttttactgg cttcagtgaa 29220  
 aaacctattt gcaaaaatagg tattaccatt gttaataaac ttagcaatta cgatatcgcc 29280  
 attttttact ttctctccat cactataagg agtgtcataa attacatgaa attcgtcacc 29340  
 ctctctaata tcaaaaacga agtctacatc ccagccaaaag atataagcaa agtccataat 29400  
 cacacttttg ggaatatattg catctagagc tgcctcataa aatgaggatt gaataatccc 29460  
 gctattgtaa gattcaatta actcaatact tttgcttata ttcttaata ctattttctgg 29520  
 agtaaggcta attgaaatag aatttatttg atctttcatg atctcaattc tggttagctc 29580  
 ctccaccagaa tattcaaaga gcattttttc ccctgggtta atgttagcaa tgatattttt 29640  
 ggagtcta ctaaaaattt tataagcagt atttaaaggc actgaaaagt tttcaaaaaat 29700  
 tattgaaaga ttttctccat cttgcacctc atgcatctga tatgtttttcg tttgctcgag 29760  
 aggtatagag aattttttcag taatctcaat ttcttcaact ggcaaaagatt catatgtttc 29820  
 tatatcaata taaagcatta atattaagac aatagaaaaca gcaaaaagaaa caagtaccgc 29880  
 tcttttttgg acttttttaa aacctatcat tagtcctctt gaatatattaa gagttcagtg 29940

09/847,513 DeLong et al.

ttgccacca atgtaccga gttcttttgag acaacctttt cgtaaagaaa ttgcagaagg 30000  
 taattaggcc caccagcctt ataaccgctg cggatagat tttgtccacc aaaaggctga 30060  
 gaaccaacaa cagctccccc catatctcta tttatgtaaa tatttccaac attacacttg 30120  
 tcactaaata tgtcggccct ttttttctact ctagtatgaa taccatttgt aaggccgaag 30180  
 ccactatcat taatattttc aattagtgc tccattttcat ttgattttaaa tctaacgata 30240  
 tgaagtattg gtccaaattg ctcgctctttt agatcagaaa tattatcaat ctcaattatt 30300  
 gttggagcta caaaattttc atctacatta tcagcagacc taaaaataga ataattctta 30360  
 tcttcaaaac ctttgacgta agcattgaga ctatcaagag atgtttttgct aattattgga 30420  
 ccgatatctg tgtcaagatt ttgagggttt ccaattttta gttctttcat accaccctta 30480  
 atcattgata aaagatcgtc atatatattcg tcttgtacac agagaactct caaggcagag 30540  
 catcgctgc cagagctatc aaatgctgaa cgaataatat catcgggttg ttgctcgaga 30600  
 agtgcgcttg aatcaacaat cattgaattt atgcctcctg tctctgcaat caaaggaata 30660  
 attgattcat ggttacttgc aaggctgctt tgaatttttt ttgctgtttt taaagaccgcg 30720  
 gtgaatgcaa cccctttgat gttattaacc tttgaaagca tgtcaccatg aatgccgtct 30780  
 cctaggatta aatttaaagc atctttttggg accccaatct catgaaattt gtccacaatg 30840  
 atgtagcaa gaattgaagt atgttctgag ggcttaactg ttactttatt tccacatgca 30900  
 agggctgcac ttattgccc tataagtatt gcaacaggaa agttccatgg actaatatcat 30960  
 aaaatatgcc ctttttggtga gtaagatagg gcattgatct cgccgggtgg accctcaaga 31020

09/847,513,0000001

atatggtctt ctgtttggag gccaacagct tgttttgcac agtatctgag aaaatctata 31080  
 gcttctctaa tttcatcaat agtattttgt actgttttttc cggtttcatt cataaggtaa 31140  
 taaattagct cagaaggatt tgcttcaatg tcgtctgcaa ttttttctaa aatggatgct 31200  
 cttttttcaa catgcatcaa agaccattca ctaatatattt gttcttcaag gctgcccttt 31260  
 atataatcta gatcatcata tgaggctgta ccaatatattc ttccatctgc aagagatgag 31320  
 atgtcatgag tattgggtttt cttataatcc ttgcttttat aaatcgatga agcctttatt 31380  
 tcttttgaat caaacctttcc aagctcctct tcaagcattt ctaaaattcac cctttcacct 31440  
 aaatcaaatc ctttagagtt cgtctctgca ttaaatatat ttcttggcat tggatatctct 31500  
 tttttttcat cctcaatttt taaatggggc cctctggcaa gccaaagcaga atctgttttca 31560  
 ggatcaagta acctattaat aaatgagctg ttgctccat tttctaatag tcttcttaca 31620  
 agatagggtg gcaagtcttt atatttgcca atcgggcgcac aaattgaggt atttttttca 31680  
 gtatttaaga ttttgtttgc ggacttataa agcagctctc ccattccaaa tagtctttga 31740  
 aattcataat ccttatattga gccaaagatga tggattgcag aaatgggatg tgcattatgt 31800  
 gtagcaaaatt tagggtaaat tttttccaca ttaaaaaattc tttttgcaca ggctaaataa 31860  
 gctaagtctg taacagattt ttttgtatag acaggataac catcatagcc atagatttga 31920  
 gcatgcttaa tttcataatc ccaatatgct cctttttacga gccgaacatg catagggtgct 31980  
 ctgtttttcta gtactcttc taaccaatct attgtggcta tagctctttt gccatatgcc 32040

09847513 . 000001

tgaacagcaa taccaaatcc ttccagttt ttaatatattg gtgaaagggc catctcttta 32100  
 ataattctcta aactaacggc cagtctgtct tgcctcctcag catcgattgt aatttctaca 32160  
 tccttagact ttgcatactc tgtgagctga ataagttttg gaagtagatc agacttttatg 32220  
 tcttttagct ttttcatttc atatctagga gataaagctg atatttttat tgagacacca 32280  
 ttaattgtat tttttgtcag atttatcttg ccgacttcat ctatagcatt cttataagac 32340  
 tgataataag tgcgggcttg ttcagcgttt cttgctgcct ctccaagcat gtcaaatgaa 32400  
 taaatttcat tttctatatatt cttaattttt tttatgtcat caaaatctct gcccataaca 32460  
 aattcttgac tgagaatatg catggcacca actacagcat ttctaattgg aaactcgcca 32520  
 gatttttgaga ttaagaact taataggcg cttgggtttt tggccactc atcgggtgta 32580  
 gaaactacct tgcctgcaag aagaagacct catgttgatg cattaacgaa aacactatcc 32640  
 gctttgttta aatgctctat ccacgcacct tcagataact tctcagatat tattagatcc 32700  
 cttgtttttt tgtctgggat tcttagtatt gattcggcta aacacatcag agcaacacct 32760  
 tctttattat ctagcccata ttcgcttaga aatgcatcta gttttgttct ctcgctctta 32820  
 ttctctctgc aagcatcaat tattttattg agcatttttt gaaatcgaag ggtcatttta 32880  
 aaaatcagaa ttactaataa gatctgaaac tatttcttgt tcaggataaa acttgttact 32940  
 tgttaatgtc atagctatta ttttaattct atagtgccaa gtatcaactt actgattcta 33000  
 ttaatgactt aaaaacatta tgatagtcac atgagtgatg cacttaaatt aattaaacga 33060  
 ggaaccgacg agatcctcac agagtctgat ttaaaaaaga aattagattc tggaaaacag 33120

09/847,513 : 09/847,513

ctaatcatta aggcgggggtt tgatccaaca gctccagacc ttcatttggg tcataccgta 33180  
 ttattaaata agctaagaca ttttcaagat ctgggcata aagtaatttt tctaattgggt 33240  
 gattttactg gtcaaatagg agatcccctcc ggtaaaaata aaactaggcc aacactttact 33300  
 tcagaagaat taatttcaaa tgcaaaaaaca tatgagaaac aagtttttaa aattctttaa 33360  
 aaagaattaa cagaagttaa atttaattct gagtggtgca acaagcttgg tgcagatgggt 33420  
 ttgattggtc ttgcatcaaa atataatggt gcaagaatgc tggagcgtga tgattttta 33480  
 aaacgttata gcgcaaatca aagcatagct attcatgaat tttataccc ccttgttcaa 33540  
 ggatatgact ctgtagccct agaggctgat gtcgaatgcg gaggaacaga tcaaaaaattt 33600  
 aatttgctag tagggagaga gctacaaaaga tcctatggtc aagaacctca ggttggttta 33660  
 actgtaccca ttctagaagg cttggacgga ataaataaaa tgtccaaatc attaaataac 33720  
 tttatagcaa tagatgaaga gcctaatagat atgttcggta aaataatgtc tatttcagac 33780  
 gagttaatgt ggagatgggt tgagttactc agctttacct cagagaaaaga aataggagtt 33840  
 cttaagaaga aaatggaaga agggaccac ccaagagata ttaagtcttct tttagcagaa 33900  
 gagtttagtag atagattcca ttcagaggggt gatggttcga aatgcaagga agcttttctt 33960  
 caaagatttc aaaaaggta aatgcctgat gacattcctt ccatgtcagt tgatgttggg 34020  
 gctgagggca ttccattagt aaacctttta aagaattgtg agatgacatc aagtacatct 34080  
 gaggcgatga gacttgtaa acaaggcgga gttaaaaatcg actctgtaaa aatagaagat 34140

09/847,513 DeLong et al. Corrected Sequence Listing (August 4<sup>th</sup>, 2001) 34/235



cctaaatgc taatttcaaa aggccaggag tctattttatc aggttggtgcaa aagaaaaattt 34200  
ttaaaaaatta aacataaatg aaaaataaat tagtccaaat atttcttctt ctttttatag 34260  
ttgcgtgtaa tcaggatagc cctaataatta aaaaaattag taatatgcag tatttttattg 34320  
ataatgaaat aagggaaagga atctattctg ttgaaccagg cttgcaatac tcaattatcc 34380  
aaaatggaga tcaaagtctt gaatcaccat tgctgcagga tacaattaca gctcattttc 34440  
acgggaccct cactgatggc tcagtttttt ggagttctgt tgaaatgggt gagcccttaa 34500  
cagtcgaact atcaggtcta atagttgggt gccaaaaaat aatctctatg atgaagaaaaag 34560  
gtgatgaatg gagagtttat atcgacccaa gtatggccta tggcgatgag ggccggcctg 34620  
ggataccttc aaactcaatt cttatttttt atattgagtt attagatata caaaaaaact 34680  
aacccttata tataacggat agagcgtatc cataaacctc agcaacctgg ttaataaattt 34740  
tatttttagg tgttcgggt ccatgaccag ctcttccttc aattctaatt aatattggat 34800  
tatcacacc ctgggactct tgaagttttg ctgcaaaact gaatgaatga gaggtacta 34860  
ctctatcatc tcttttagct gtagtaatca gagtagttgg atagcactca ctttcaacaa 34920  
tattatgcag aggcgaataa gctaataagt ttccaactc atctttctta tctggagagc 34980  
catagtcaact ttcccaggcc cagcctatag taaatttatg aaatctaagc atgtccaaga 35040  
cgccaacttg tggaattgcg actttaataa aattaggatt ttgcaacatg gtcgcagcca 35100  
ctagcaaac tccattagag ccaccttgaa tagcgggtga tgatggagaa ccaatttctt 35160  
gtgcgtgtaa aaatttagca gaatatgcaa agtcatcaaa aacattttgt ttattaaaaa 35220

09847513 DeLong et al.

gcctgcctgc atcatgccag ttatcaccat attcaccacc acctctttaga ttaactacag 35280  
caactattcc cccttgggtc atccaagtga ggtagctttt actaaagcca ggtagtcttg 35340  
agatatataa gccaccatat ccataataaa gtattggagt attactgtct atttttaagg 35400  
actttttgta actaagatgg atgggtattt gtgttccatc ttgggatgga aagaatttaa 35460  
agtcagatgt aaataaaagt gagtcgtgtc cttaagatc ttctttccaa aaaagtctct 35520  
gtgacatatc agttagatta attttgtata ttctcggagg agttacaaaa ttgttaaatg 35580  
aaaaataaga tacctcatct tcaatttcac caccaaaacc acccattgtt ccttttcttt 35640  
ctgttgctag cttattttta tatgtctcct ttaaatcaaa gaagtgaacc tcagtaaaaag 35700  
tatcaacgag ataagaaact acaatagaat tatttataa gtaaacacta ctaatagaat 35760  
tagtactttc accaacgact tcattccaaa caaatgagcc attttttatt gtaagtata 35820  
ctactttttc gttcgcagca ttttcagttg agtaaaacca aaaagtatca ttcttgcttt 35880  
ctaaaaagct ataagcgct attagctcat ctattaaagg aataaaaggc tgatcagggc 35940  
ttagctgaac atagagtcta tttctttcat ctgtgccttc accaatagat agaaacttaa 36000  
ttttagaatc ttttacaaca cttattcccc aactccatct tggcttttca ggattctcat 36060  
aaacaatcac atcctcatct tgagcagtac caattttatg aaacattaac ttgggagctg 36120  
tattaatatc ttttaaaagc tcttcgatg gctcgtcgtg tttttggtta taaaatccag 36180  
aatcatcatt ttcccatgaa gcaccagaaa atttagccca ttcaatccta tcatcaagag 36240

U9847513.0000001

tttttcctga ttcaatatca agtactttcc aggttctcca atcagaccca ccatctgata 36300  
 ttgagaaagc tagaagagat gcatcattac taacgctggt acttgcaaga gagatagttc 36360  
 catctttctga gaactgattt ggatcaagta aaactctatc ttggcattct tcacagtcct 36420  
 taatcatcag cttgctttgt tgccatgacc catcattgaa ataataaaaa gtttttttgt 36480  
 taacctgata aggcattgctt atcgaatctg tatcccaaac ctcatctaaa ttcttagcaa 36540  
 tagattttttt atattttattt tggcctataa atttttgtgt gaaattattt tgtctctcga 36600  
 cccagtctgt tgagtcctca cttgtaaaaat cctccatcca tctataagca tcttcgatta 36660  
 ggtatccatg gacttcttca ctaaaaggaa ctttatttga ttcagggtat tcaaaaattat 36720  
 ctgttttgatt tgagcaacta actagaatta atagagctgc catgtatatg tattttattgc 36780  
 gcataaaaaag cataatactt tacaattacc ttatgaatca attcaatcga gcatggcaac 36840  
 ttttacgcaa aaaattggaa gtttttttatt gtacttgcg caccatcat gatgcttgaa 36900  
 atagcaatgg catcaatggt aaccccaata caaaacgtta ctcagcctga ggatattcta 36960  
 gaatttttca atgagaatat agccttttcta ggctctgtta gtctcttagg cgtagtttta 37020  
 agtatggcctt ttatgggagc gcttttttgtt tcgtatgcac ccatagaatc agaaaaatgag 37080  
 attgagcctc taaatgcctt attttttaggc ataagaaaaat tctttccact tctgggggct 37140  
 tatctttatag cttcagttgg tgtctttttt ggtatcttat tattaatact cccagctttt 37200  
 tatgtagcag caaggctttg catttttctt gcatttatta tgcttgaaga caaaggagct 37260  
 atagaatctc ttaaattatc ttgggaaaaa acggatgagc atggcaccac tttgtttggg 37320

09847513 .0000001

cttaccatta ccttcttttc ttttaacaatg atttttgcac cagttgcccc atccattata 37380  
 agtccgggat taatgcaatt agttgttctt gcaattattg aatatgtaat agtaattcca 37440  
 tggggctatg tataatttttag tttatacaag tcattaaaaa gatattaggc agattaacta 37500  
 gaataaatta atacgaagct tttcgttttac aaaaaaacac accactctta taatccttcc 37560  
 tccgaagatt tgcactttgt taattcatgc aaacaacttt gggctctgtag ctcagcttgg 37620  
 ttagagcgca cccctgataa gggtaggtc ggtggttoga gtccaccacg accaccatt 37680  
 ctttgaacat taactttttat ttcctttata cttattctta actaaataag gatagtatta 37740  
 tggctaaagg tttagacaag caaaaaaacg acaagaagaa aggcaaaact ctaaaaagaaa 37800  
 aaagagcggc aaaaaagaa aaaaaagaat agttttttgt agaagactat tattatctag 37860  
 tcgacctagg gccatcttta aaatctatct taaagaacaa cattattccg gtcactgagc 37920  
 ttactaaagc taaaatagaa aatattttta aaagaagggt gtcgatatta tctcttttctt 37980  
 tccaatccat aatatgaaac ccccatca gatcccatat ttccatttta ttagatctaa 38040  
 tggccacaac ctctgctgag taaacattaa tgtaaacatt taattcagta cctttaacat 38100  
 ttttactttt tactctataa attggcaatg atctgcctct atattcagat cctgattttt 38160  
 catttacaac ttcttcggtg gcaaatggca aaagagttgt ttggcttgaa actgaatcta 38220  
 tagcatcttc catagaaaac tttgtttaaag gctgccccaa catattcaga tacttagttg 38280  
 agccctttgt tgtgatgata acaatctctt gacctaatct ttttttgaat ttaacttctt 38340

09847513\_080901

tagctttctc tatttcaaaa tttaacttac ttaaatcaaa agatgtttca acatgattaa 38400  
ggtattgctc gcctcttacy agctcaatct tattaaaagc aaagtatatg cctgagatag 38460  
tccaaagtaa taactgaaga gagataaaaa aacttaggta cttatgaatt ttcttaacta 38520  
aaaagtccat cttttcataa tcttcattat ttcatcaatt ttacatcaa attcgtcatc 38580  
accttttaaa gattcccttaa tgcattcacg caaatgggtt tttagaattt tgcctcaac 38640  
agtaataaatt gaattcttaa gtgcttttat ttgattaaga atatcaacgc agtattgacc 38700  
ttcctcaacc attccttgaa cgcctcgcac ctggccctct attcttttta agcttattaa 38760  
ttgctcttta tgacatggat gtttcatttt ataaaaacct cggaactata tacgtaaaaa 38820  
ttgtagcgaa taaaaaaatt gctacagatg catagtacag atacctagat ttttttctag 38880  
tttctgaaca tgctctgcct agttctgggt cagcagggca aggcatatga taagttttgt 38940  
agttgatata tcctgcaatg acaatcatta ctaatgcaaa tagtgttatg taaagcttat 39000  
attgagatag ggtaattaag aatggaaata cagttaccag gcttgcaaaa ctgccccag 39060  
caccaagagc acaaaaaata gctggcaagg cacaacatat caagtagaa gaagatgcaa 39120  
ataatgaaa aaaattcgac gccttatcat tcatagttat taaatttcta ggatttgaag 39180  
atcagacata ttttggccat tagaaaagaat cttcttttcg atctcagaga aggtgatttc 39240  
ttttgcctta gaaaaggcaa tcactacttt tccattatca agatcaacat ctattctttt 39300  
gacatcttta tcttttagaa aagtttttga aatacctctt gcacagaaat cacaaaccat 39360  
tcccttcacg cttacgatcg ctatatttac atttgagagt ccttcaacaa atgcatcaaa 39420

9947513 000001

cctctcaggg ttacatcaa gatctttttcc atcaaccatg gtttcatgta aatgaccttc 39480  
atgagaatgc ccatgcatac cactatggtc attctccgcc gcaataaggc cgcttgaaaag 39540  
caaagtaatt aataataaatt ttttcatatt tttctccttt taaaatctgt acataaaaatg 39600  
taaatcccg ctcgtatctg cgtcatagcc aaactcaaca agggcattac cataaaaaaaa 39660  
tttcaattct ggatacgtag atgatttatt tgttacagta ttttttttgg ttttaaccat 39720  
aaccatgta tgtaaatcat tgtagtcagc aacgtaagga gcaaatccaa gttgaatata 39780  
ttcttcttta tagtctttgc taaatttact gttaatatcc ttttaaccaa agccgcata 39840  
ccacttccgg gtttcccagt ctccatggat accataaaaa ttatttttaa tacttccagg 39900  
cacaattcct gattgaaaat ataaattcct ttgggatgta aaagtatttt ttctattcac 39960  
caaataagta aaacgaaaat agttttctgt gttcttataa aatttatttt tagccctttc 40020  
taggccaatg gaatatattgt aggttgga gtagtgga taaatagaat cattaagggt 40080  
atctgattta tacattatgg ttgttccgcc agagtaggat atcgggtcttg catcagcatt 40140  
aaagaccatt aaaaaaacta aaatgattat ttatatata cccatggggg gtatcttacc 40200  
ataggattat aaattaattt caggcatctc tccaagcgc cattttcaat gaaatcacca 40260  
tcagaattcc aaaaataaacc atagctatag cagatgcata aaaaatacta actagcccat 40320  
atgagagcca gtaaacagct accaatgcta taacagaaaa tactataagc ctgatggccg 40380  
tgattataat tggccatttc attgcattgg ctccctgaga tgcaaaatac agggaaagcc 40440

09/847,513 DeLong et al.

catatccttg aaatacataa catacaccaa ggatttgaat atactgcttt gtaactaata 40500  
aggtctctgg atccgatgta aaaatactaa tccaaagatt tgggtttagt gctaaagcta 40560  
gaccgataac agccgaaagc aatccagctg ttgttgcacc aaccatgcct attttttctg 40620  
ctcgcctctat atttttggca cctatatattg ttccaacctat ggctgtcata gcagtcccaa 40680  
tgccgaacac aattgggatt aataagaact ctacccttga accaattcca taaccggcaa 40740  
tagcagacgt tccaaattga ccaatgagtc ctgtcagtaa aagaactgtt gcaactgtca 40800  
ttattggcga aagagatgcg ggtaaggcaa cagaaaaaat atcttcgaat aattctcttt 40860  
caattgttag tcttttgagc ttaagcctta caggcgatga agggctcgaa atttttatta 40920  
aagttactaa tgccataaaa cctgaagtaa ctatcattga ccaagcagat cccacaagcc 40980  
caagtttagg tagaccgaat gatcctaata taaatccagc agagaaaaaa acttgtatatac 41040  
ctgcacaaat taccgttagg actgcaggaa attgcatatc tccatccct cttagggcgg 41100  
cggttaagct acctgatagc caaacaacta ttgctcccaa taaatagaca aaacaataag 41160  
ctaaagactc ctctaataat gcacctgttc cacctaatat tcttaataga ggctctccaa 41220  
aaataaaaaa gataattaa aaagctaaag ctccaaagca agatatatac aaagaatgcc 41280  
ataataattt ttcagcccta ggcttgtctg cggctcccaag acttcttgca atagaagatg 41340  
tgaccgctcc tccaagagca ccaaatgcca tctgtctgagt cagcatgatt gcaggaaaatg 41400  
ccaaagtaac agcagctaataat ggagttatgc caagttgact aataaaccag aattcagcaa 41460  
gaacaaccac agcattttatt aaaaatgcaa cagtattagg tgctgacatt ttaatcagca 41520

09/847,513 DeLong et al.

aaggaaatat tggatctttt aaaaaattgct ctgtttctctt gtccataacc ttctcattta 41580  
ttaataattta ttgtgtttaa gggtgaaaga gaacacgttc tcaggattca cgaatcttta 41640  
aataataata acagctttta gttaatttaa tttttaaaaa caaccaaatac tattatctgc 41700  
tagattttat tttttaaaca tattaatatc taatttaatt gcaaaggggt ggccatttcg 41760  
gcctgagatc cttggaaact ctttgaacct gatccataca ataatggcgg aggaattgca 41820  
tgaaaaatcat taaaaatcat ttattttatt caataatctt actatatattt tttagtctag 41880  
agacgagctc tcaagctata gaagaagtta ttatcaaagg agactggaga gaaactagtc 41940  
tgtcagcaga agactcaagc attgcagtc ttgatagtaa attaatagaa tcccaggctt 42000  
taaagcactt tgaaaacctt tcgtacctag taccaaaattt aaattttgct gcaagtgtt 42060  
ctagagcaag acatttccag ataagaggaa ttggagaaaag atctggctat gaaagaactc 42120  
caaactctgc agtaggttta ctgattgatg atatagactt ttctgggcaa ggtggaattg 42180  
ccacaacttt tgatgttgat caaattgaag ttcacagagg cccccaagga gcaagaatag 42240  
ggtctagcgc aatggcgggg ctcatatata ttctacaaa agacccaaca gaaagctttg 42300  
aagggaaagg tgaaatagtt atggggtcac atggaacttt taataccgga attgctgttg 42360  
gcggtcctgt aaattttaat aaagacctca cttataggtt agctataaag aaagattatt 42420  
ctgacggatt taggaagaat atttttctca acaagtctga tacttctaag aaagatgaga 42480  
gcacatttag attaaaagt aattgggtaa ctgataacca aaccacttat aagtttctta 42540

09847513 DeLong et al.



tatctcaaat agagttagat gatcctgcag atatatggac tatagatggg agtcttaata 42600  
 cattatctga taagcctgga atggactctc aaaaaagtaa tgcctatggg gtaaaaaattt 42660  
 atcatcaatt taaaaaattt gaatttcaaa gcctatcaag cttaacgaat actgatgtca 42720  
 ttcttagtta tgatgctgat tggggcaacc ctgagtcgca ttcacccttt atttatgact 42780  
 atttttcaga aactacaaga aaaagagata ctttttagtca agaattttaga cttgtatccc 42840  
 aatttgcaga taaaaatata gaaaaaagca tcgaatgggt cgttgggggt agttttgtag 42900  
 atataaatga aacaaatgct aaaaaagata ctggtattta tggagatcca tcagacccat 42960  
 atggtcctta ttttagtaac tcttcttctt tgagtgaactt ttcttcttca agctattctt 43020  
 tatttggaaa tattgattat ttaattaatg aaacaataaa aatttcaatt ggtggaagat 43080  
 gggagaattt taaatctaata tatttttgatt cttatgacga atcatttttca ccatcaata 43140  
 aaatgtctgg tggtaagtgg tcactagtta aaactcttaa taataattct aatattttatt 43200  
 ttaatatatgc taagggctat aaccagggtg gatttaattt aggtcttggt cttgataaaa 43260  
 attcatcaaa tagaaaattta tattatgac cagaattttt aactaattac gaagttggaa 43320  
 ttaatagtaa atttttccag tcaaaaattaa atcttggagc agtcctgttt tattctgac 43380  
 gaaaagatca acaggtctta atttcaaccc aggttgatcc ttcagatccc aacacttttt 43440  
 tatacttaac ccaaaatgct gcagaaggaa ttaataatgg cttagagtta aatatagatt 43500  
 atgcactgaa taaatctcta ggtatatattg ctaattttgg attgctaaat acagaaaataa 43560  
 aaaattggat ttcaagacca gatatagaag gtagagaaca ggcgcatgct ccaaaaaata 43620

09847513 DeLong et al.

gtttttcaat aggcataat tggaagcaa caaaccaatc ttatttatca ttgaatgttg 43680  
 ttggtaaaag tgagttctat tactctgatt ccataacaa tacttctgag tcatacaact 43740  
 taacaaacat taattatgga tatgaacatg gacaatggac ttattcatta tgggcaagaa 43800  
 atatttttga taagtactat tcagtaagag gtttttactt tggtaacgag gcacctgatt 43860  
 ttatagatac gctctacaga agacatggag atccaaggca tataggagtc atggtccaat 43920  
 atgatttcta acttttataag tgaatttttg atggaatatag cagctgttgt atttgctatt 43980  
 atttatttat tacttgctgt aaaacaagac gtaaagtgtt ggtttgcggc tataatttagc 44040  
 tcaatatttat atttctttat tatgtatgac gcaggcttat atatggaagc ttacctgcaa 44100  
 attttttata tcatgatggc ctttttatggg cttcaacagt ggagaagtgt cgatactgat 44160  
 gctcctcaat ttattgttag aacatgggat aaaaggatgc atattaagat cattacatta 44220  
 atagttatca tgactttaat ctctggattt ttattagaaa aatataccaa tgctatcttg 44280  
 cctttcatag atggactaac aacatgggga gcaatagtgt cgacatatat ggttgccaaa 44340  
 agacttcttg aaaaattggat atattggttt gttatagatt ttatttctat ctttttattc 44400  
 atgtccagag gactttttatt aacctcagga ttattcttta ttatcttgt aataatatat 44460  
 tttgggttaca tgtcatgggt aaaaaataaga gacgatatta gtgcagaatc ttcataacaa 44520  
 tctagatata aaaaatcata atctagagat tattaaaaca attaaatcag gccctgtatc 44580  
 tgaaatatct atttghtaatt ttgataacat caaagcaatc ttaagagttg atcatccatg 44640

09847513.000001

cgcacacaaa ataaatgtag atcgagaaaa tgaaattttt atactcagcc aactaaaaat 44700  
 tttagatttt agtccagagg ttttatttag tgatttgtct tatggaattt tagtatggag 44760  
 atatatcgag ggtattgaat ttccacttgg taaggattcc aatgaagttt ttttaaaaaa 44820  
 acttggaacc gaattaaaaa aaattcatga tattgatctc cccaaaaagta agaaaaaata 44880  
 ttttagcaat gacataaatt tttacagaaa tttactaaaag gaggttcctg aaaatataat 44940  
 tcttcataga ggatttgatt tatacgacaa actcaataac tctgataaatt atgttctctc 45000  
 tcataatgac ttgaataaga caaaccttct ttggagggat aggttatttt ttttagactg 45060  
 ggagtactcg agttttaaca atcctttttt tgatattgcc tcgttatcga atgcttataa 45120  
 cttatcaaaa gttgataggg caatttttat gaaagcttat acaaaataatg aatattcagt 45180  
 attaaatgat acaaatctta gagaatggat gcatttttgt cattatttag agtacatgtg 45240  
 gagtatttcg ctcatacaaa atggaaaaat tgatcagaac accctaaatt taaaaaaatt 45300  
 agagaaaaaa ttaaaaaata ttatttaaac aacaagtatt gcgcattctt tagttgttgt 45360  
 tttactatta tgtaagaaat ttacatttaa atgggggttt gatatgaaaa ttttatgtgt 45420  
 cttatatgat gatccaaaaa caggtatgcc agaaaggat gcaagagatg atttaccaaa 45480  
 gttagataag tatcctgatg gaatgacact tccatcccca aaatctatag attttactcc 45540  
 tggtagtgta cttgggttgt tatctggaga actagggcct cgaaagtctc ttgaagatgc 45600  
 tggccataca ctagtgtgta cttctgataa ggatggagat ggatgtgagg ctgataaaga 45660  
 attagtagat gctgatattg ttatatcaca acctttcttc ccataattatt taacaagaga 45720

456947513 1000000

caagatgaaa acagcgccta atttaaaaat ggcaattaca gctggcattg ggtctgatca 45780  
tggtgatctt caggcagcca tggataatag cgtggatgtt gttgaagtta cttactgtaa 45840  
ttctcgttca gtcgccgaac acattgtgat gatgaccta tcgatggttc gtgattatca 45900  
taccacaacat agaattgtaa aagagggagg atggaatata gctgatgctg tacaaaaggtc 45960  
atatgatga gagggcatgc atgttggtac tgtagctgca ggtagaattg gtattgatata 46020  
gctaagaaaa atgaaaccct ttgatgtgca ttgcatcac ttgatattc ataaactctc 46080  
tgatgaaata gaagcagaac taaacctcac ctatcatgat tctgtagagt cattagttgc 46140  
tgtgtgtgat gtagttaata ttagtgtccc attgcacatc aaaactgagc acttattcga 46200  
tgatgaaatg attagtaaaa tgaaaagagg tgcatatatc atcaatactg ctcgtggcaa 46260  
gatttgtgat aaagatgcta ttgcaagagg cttagagtca ggcagctaa gtggttatgc 46320  
tggtgatgtt tggttccac aaccagctcc aaatgatcac gtatggagaa caatgcctaa 46380  
ccacggcatg actcctcata cttcaggtag ttcgctatct gctcaaaaca gatatgccgc 46440  
tggggttaga gaaattctag aatgttatct tgcaggcgaa ccaattagag acccatattt 46500  
gattgttcaa aatggtgatc ttgcagggtat ggggtgcgcac tcgtacacaa agggtagacg 46560  
cacagatggc tcagaagagg ccgctaagta taaaaataa gttttagaac ttacttagcc 46620  
tttccttaat aatagagtct gcctctgcca taatgctatg cattaattct tccactgtcg 46680  
ggatgtcatt aaccagtcca gcaaccattc cacatgacca ggctccaacc tccatagttc 46740

09/847,513,000001

ctctcatgcat aatthttttgga tagactctctg caacctcatc cacaatatca gcaaaagtta 46800  
 attcatcacc aagagcttttt tctthttttcaa tcaatctctc aacagcttca ttattaagaa 46860  
 ccctthtctgt atthgtttaat gatctcatga tgagtctagt atctaactca gaagcatttta 46920  
 cgatagcctc thtcacatttt tgatgaacag gtgcattcttg agtagcaata aacctagtthc 46980  
 ccataattcat tccctcagca cctagtgaac tggcagcaac caaactthctt ccattctgcca 47040  
 thccaccgga ggcaacaaat ggtattthcaa gctcatctgc tgctctaggt aagagtatga 47100  
 aattaggaat gtcatctthct cctgggtgtc cgccacactc aaaaccatct acagaaaaccg 47160  
 catcgcaacc tattgcttgtt gctthtttaatg agtgccttac agagggtgcat ttatgaataa 47220  
 cthtttattcc tgctthctthtt aaagctggga ggtattctgc aggtattthctt ccagcagthct 47280  
 ctacaacagg aaccctctgca tcaataatta cthtttatcaa accaggatag tctgggggtg 47340  
 thagtgtagg taaaaatgth aaattaacag caaatggctt attggtcatt tctthtgcattc 47400  
 tggcaattthc atthtgctaat thctcaggcg taccctgtgt tagaccagta attgtthccaa 47460  
 gtccaccgca atthtgatact gccgctgcaa gctcgggcaaa accaacatgg tgcattccac 47520  
 cthtgaatgat tggatgtthct ataccaaaata attcagthtat thtagthththc ataatactthc 47580  
 cctattthtattc thccatttggg gthtaaaaaat cthcatattg agthtgaagt cthtgcattgc 47640  
 cacttatcca thcatctcttg thattacctt thctththtaac atattthtaga acctcactat 47700  
 gaggggttac taaaaacctc thctctctcta tggcgthtaag aacatctthth gcaacaatat 47760  
 cagctthccat cattccatca actcctgcta caccaggacc atttgcagtc atagcagthc 47820

09847513 DeLong et al.

taactgcttg agggcataaa caggaacgc cgataccttt atttccatat gtaattttta 47880  
 tccattcagc aaaactcaca gctgcagctt ttgtaaactgc atatccagcc gcacctagtt 47940  
 gagttaaaag accggctgct gaggatgtat tcataagata gccttcacct tgttctatca 48000  
 tttgaggaag cacatgtttt gcagcatgaa tatgagactg aacattttaca ccccatatca 48060  
 tgtccaatc cgaagtatct gcttcaaaaa atcctgggctt tccaccgata cctgcatttg 48120  
 aacaaaaat atcaataacca ccagaaaatt cattagcctt ttgtataaca ttataaatgt 48180  
 cgttttcttt ggaaacatca gcacttactg caagcccatt aacactcttg gctgtttcct 48240  
 cgcgccatt taaattcata tcaacacata ctatagaact tgcccagat gcatagaaact 48300  
 cttcacataa agccttacca attccactgg cagcccctgt aaccacaaact cttttattat 48360  
 taattttcat aaattgacca ctttttttaatt tttttattta actttattgc ttagatagtt 48420  
 tattacaaa caacaatact taatattgca aaatcgcata cccatcttta taaattttatt 48480  
 gctattatta gtgctagaaa atgaaatact gagatttaat atgaagaatg ttgttggttat 48540  
 tggctcgtcc ggggcaatag gaaaagcctt tattgatagc tatatcaaaag atgatgatgt 48600  
 tgaaaatata ttttcatttt caagaacagg cctttccatt gaggataaaa aactccatag 48660  
 ttttttttatt gatattgagg atgaaactag tattttgtgat gccgcagaga agatagacaa 48720  
 gtcctcaata gatgaaatta tcgtcgcaag tggaataactt cataataaaag attttgggcc 48780  
 agaaaaaagt attagagatt taaatgcaga taacctttta aaggtcatta aggttaatac 48840

09847513.0000001

tatcggccca	acaattgttg	gaaagtattt	cattccattg	ctaaataaaa	aagaaaaaag	48900
cgtcttagca	tttttaagt	caagagtcgg	cagcatttct	gataataaaa	caggtggttg	48960
gtatgcctat	agagcgagta	aaactgcact	taatcaaatc	attaaaaagt	ttagtattga	49020
attacgaaga	accaatccaa	atgccattat	ttttggtcta	cagccaggaa	cagtagatag	49080
tgaattaagc	gaacctttta	aaagaaaatgt	aaaagaaggt	aattttattta	ctccagaata	49140
tagtgtattg	cagctaaaaa	atattattga	tacagcaagt	ccatctgatt	caggcaaaact	49200
aatttcttgg	gatggggaag	agattcagcc	atagttgggt	atgaatatat	tttcatatta	49260
gaataaaattt	ttaggggaaa	aaatgagtat	taaatattat	gactgggtcta	aatttccaggc	49320
caacactagg	ccaaataaaag	ttgccataag	agagctagat	aataacaaga	tctatactta	49380
cggagaattg	gataaaaagat	catcaaggct	tgcatacat	ctccaaaagt	caggaataaaa	49440
aaaaggagat	cgtattgcga	tactatcgct	aaattgttca	gaatttttttg	agctagaatt	49500
tgcttgcgga	aagattgggg	caatagagat	accattaaat	tggagattaa	caaaaaccga	49560
gctcagttat	attcttaatg	atagtgagcc	aaaaactcta	atttatgaca	atcagtttga	49620
agaaatggtg	aaagagctaa	aagaagaatg	taatatttct	gaaatcatag	ctcttgatca	49680
atttgaccaa	gaaagtgatt	atgaaaaagt	tttgagtaat	gcttcaggca	tttattatca	49740
ggaagaagtt	gatctagaag	ataacattat	gattatgtat	acctctggaa	caaccggtca	49800
ccctaagggc	gccatgatca	cgcacaaaat	gcagcttttt	aatgttatta	atttaggtat	49860
ttcagcagct	gtttccccctg	aatcagtcca	tttagttgtc	cttcctttat	ttcatcacagg	49920

cggaatgaat tgttattcaa atccaattct tcatgcaggt ggcgagttaa tattacttaa 49980  
 agagtttgag cctgggaaag ttctatcaat tatcggcagc tctgactatg gagttactca 50040  
 tctgttttgca gttccagccc cttatcaatt tatgatgaat catccagatt ttgaatcaac 50100  
 aaatttatca ggagttaagt atgctggagt tgggggcgca ccttgtgcag aggctatttt 50160  
 gaagacttat ataagcaagg gtgtttcgat gcagcaagga tggggtatga cagaaactag 50220  
 tccagggtgct actggtcttg aatcgtccga ggctgaaaga aaaataggat ctgctggaaa 50280  
 accgcttctt cactactgag tcaagggtggt tggagatgat gggaatgaac tgcctgctgg 50340  
 agaagtaggc gagatttata ttaaaggccc aaatattaca cctggctatt ggaagaaaga 50400  
 agaggctact agagatttctt ttgaagacgg gtggttaaaa acagggtgat ctgcttactt 50460  
 tgatgacgag ggtttttttat acatagttga tcgatggaag gatatgtata tctcagggtg 50520  
 agaaaatggt tatccagctg aagttsaana tggtatctat cagttaccac aaatcgaga 50580  
 agttggagtt attggtattg atagccctaa gtggggtgag actggtaaag cctttgttgc 50640  
 tttaaagccc gatcatgaat tgacggcaga agaagtcata gatcattggt taaaaaatct 50700  
 agcaaagtac aaaattccag agaaagtga gttattgca gctcttccaa gaaatgctac 50760  
 aggtaaagtt ttaaaaaagaa cattaagaga tatgtaatat ttttaagcca aaaaaaccc 50820  
 agcctaagct ggggtttttta tttaactatt taaagttatc taatgttaag cattagaaga 50880  
 ttctttaaca gcaacattcc atataattaa accaaataga atcttgttaa caaagtcagc 50940



aaggttatag ataaggttta agttaagagc tgatccaccg tcaccocatca ggtaacctgt 51000  
gaaataacct acaggataaaa tcgcccacc aaagatgata atatacatca ttgtgttgta 51060  
agctgattgc acagcaggac ttgcagtatt acatgcagat tttcctttctc cagcccataa 51120  
ttcataaaatc atgtataccc aagctaaaaa cccaataatg aatgcaggcc atgcagccat 51180  
gattcctgct tcaccatgt aaccaaaccac aagcataaca agagaaccaa ctagtaattt 51240  
cttaataaat gatccagcaa cattagtgc agcagcaaga attaagtaga attcacatat 51300  
taatagagga actgttagta accaatcaat gtatctaaat acagttggcg aatcaccagt 51360  
ttcaatccat acccctctca tgtacatgta atgccagaaa gcaataccag taacaagacc 51420  
agatacagtt aatgatgttt tccattttgc agaaactcta tctctttcaa caaagaaaaa 51480  
tacagtagat gctaataaag cagcagtaac taaccaaaaa gaaacaccag tgtaatcact 51540  
agcatcaagg tcaccaccac ctgcagcaaa tgtaggaagt gcaataaacac tacctaatat 51600  
cagtaataat ttcatatata actcctatat taatgacaaa gaagcataat tgctctcccc 51660  
ctaattaaat atgaaattta atagtaacca atattaatag aggttacaaa taaaaagcat 51720  
ctattttttt aatagaaaaat atgtgtaata aaacaccata ctttaaatta atatttattt 51780  
tgcataagta gacttgttta gtatcatatt taacatgaaa gttgcaattt atcctgggtc 51840  
ctttgacccc atcacaaatg gtcacacgga tattattgat aggggttgcg gactctttga 51900  
caaggttggt gttgcaatag ctaagagtga atcgaaaaac cctcttttta gcctagagga 51960  
tagaattaat ttagcccaat ctatttttaa aggaaatgaa aaagtagagg ttgttggttt 52020

09847513 DeLong et al.

tccaagaaag ttaacagttg atcttgcaaa agactatgga gcttgtgcaa ttataagagg 52080  
cctacgagca gtttctgatt ttgaatatga atttcagtta gcaacaatga ataggtcgct 52140  
ggctcccaat attgaaaagca tttttttaac accaaaaagaa agtctcattt atgtatcttc 52200  
tagcttaatt aaagaaaatat cagacttaaa aggcgatata tcgaagtttg ttcacacctat 52260  
agtagagcag gcacttcgag cgagtagaca cttagctctg acaagcttca caaaaaaaag 52320  
tagctctttg attgacgacg gttttacaaa tagtgccttt gcattttattg cagggctcgc 52380  
cttctcttcc atatacatctt aattttaaatt taaaatatcc tggactacca tcagctgagt 52440  
agaagtcttt taatgtagtt cctccaaacct ctattgcttg ctctagtatc ttttttccag 52500  
cagctactaa tcttttacaa gcactctaat ctaactcatt ggcatTTTT agcggatgaa 52560  
ttttagctaa gaaaaggctt tcagacgcat aaatatggcc gataccgact acattTTTT 52620  
gattcattaa atagctttta atatttggtt tagagtgact gcacccctgaa aaaaaatcct 52680  
ttgcattaaa attttttagaa agaggttcag gcccaagggt ttttattaac ttatgtttat 52740  
cgatatcaga agtaagatgc attgatccaa acccttctggg atcattataa ataattcttt 52800  
cctcatcaaa aattaattca atatgatcat gtttgataaa gaagttttca ttattttttg 52860  
caattctaag actcccagac attcctaaat gaagaataat ttttttatca ttagatagtt 52920  
taaaaaattat atattttgct cttctttcta agcttttcgac aacttgatct tttactgagg 52980  
tttcaagact atccaccact ttccatctaa ggtttctgtt gtgtatcctt gcttctttca 53040



cctatatga aaccatgtat ctagtctaaa gagcggctga gtaaccctgg cagatgtaga 54180  
aaaagagtta tattgctgct gcagttcttt gttttagatag tattcgttcc agttagttga 54240  
tccactcaag gtaataactcg ggagaagcgc agctcttctt tgaaccttaa gctcttttctc 54300  
tgctaaatat gaatattccg ctgcttttata tgtagggtca ttctcaagcg cttcattata 54360  
gatatccaaa agacttttcag atgaaatatt aaatgagata aaaagtgcata aaatgatttt 54420  
tgtgtaaatt ttcataaactt attttaacct attaataatg tttgcagtcg aaacattttat 54480  
tttttttaatt tttattgaat ttatttttata tttcttagag tagaatatct ctaacaaagtt 54540  
caataatttt ttataataaa aaggcttttaa aaattggcta aaaattcata tgacgctcag 54600  
gcaattgaag tcttatctgg actagatcct gtcaaaaaaa gacctgggat gtatacggat 54660  
acatctaacc caaatcathtt aattcaagaa gttcttgata attcggttga tgaagctctt 54720  
tcagggttatt gttcaaacat aaaagtatct gttctaaaaag atggctttat taaggctctct 54780  
gatgatggaa ggggtatgcc aattgatgag cacccggaac ataaagtctc aggtgttgag 54840  
ctcatccttt gtaaaacttca tgctggagcg aaattctctg gagatgatta taatttctct 54900  
ggtggccttc atggtgttgg agtttctgtt gtaaatgccc tatctgatga attagaggtt 54960  
agagtaaaaa gagatttctaa agaataccaa atcactttta ataattggaga taagtcttct 55020  
gaattaaagc caattgggga agtggggctc agaaattctg ggacatcaat taaattcaaa 55080  
ccaaatccta catattttga aactatagag attcagataa aacagcctaa gcattttatta 55140

54180 54240 54300 54360 54420 54480 54540 54600 54660 54720 54780 54840 54900 54960 55020 55080 55140

aaggccaaag cagttctctg tcctggatta acgatagagt tcgttaatga aaaaaagact 55200  
gatgataaac aaaagtggta ttttgaagat gggctcaaaa gctatttgat tgattcttcc 55260  
gagggagcag acttggtttt gctagattca attgtatgct ctaaaaaatc tcaagctcaa 55320  
gagcttgaat ttgcaatcaa ttggtcatta agacccccc aaataaaact cgatgaaacc 55380  
tatgtgaatc tcataccaac tgctcagggt ggctcacatt taaatggcctt taaggctggg 55440  
cttttagatt cattaaaaga attttgtgaa tacagaaaatc tattgcctaa aggttttaaaa 55500  
attaatgcag atgatgttct taataatgca atttttataa tttcatctaa gcttcagaat 55560  
cctcaatttg cagggcaaac caaggaaaga ctagattcaa aagatcacat gtcgttcgctc 55620  
tcaagtagca caaaagacat tttaaagtatt tggtttaaca ctcatacaga agagggcgaa 55680  
agaatagcag aacttgcaat tatgtctgct cagacgagag caaaagtctc caatatagtt 55740  
gaaagaaaaga aaacttttag aggccagcc ttacctggaa aactttcggg ctgtaatatg 55800  
caggacttaa atgaaacaga gcttttttta gttgaggggg actcagccgg agggtcgcga 55860  
aaacaagcaa gagaaagatc ttccaggca atcatgcctt tgagagggaa gattttaaat 55920  
acttgggact tagaaagtgc agaaataata aaatctcagg agataaaaaa cctatcaact 55980  
gcaattgggg ttctgccagg aaataatgac ctttcatcac taagatacgg aaaaatttgt 56040  
attcttgcag atgctgattc agatgggtctg catattgcaa ctttactttg tgcattgttt 56100  
ctaaggcatt ataaatcttt agttcaagag ggaaggatat atatctcaat gcctcctcta 56160  
tatagaatcg attctggtaa agatgttcta tatgcacttg atgataaaca gcgagatgaa 56220

atagttactg aatttaaaaa gaagaagggc aagcctaaag taaacattca aaggtttaaa 56280  
 ggacttggtg aaatgaatcc acctcaacta agagagactg tgatggaccc tgctactcgt 56340  
 cagcttggtc agcttttctat cagctcaagc gataatgcaa attctatgat ggacttactt 56400  
 ttgtccaaaa agaacgcacc agcaagaaaa gaatggcctg aaaagaaaagg gtctctagca 56460  
 aaaatataaa tatgaaagaa caaataacct caattagcct caagcaatat gctgaagagt 56520  
 cttatcttaa ctatgcaatg tatgtcattt tagatagagc ttgcctaata attggagatg 56580  
 gccttaagcc tgttcaaaga agaatactct atgcaatgtc agagcttggg cttgatgctg 56640  
 gctcaaagta caaaaaatca gcaagaactg ttggagatgt tataggaaaa tttcatcccc 56700  
 atggagacgg cgctgcataat gaagctatgg ttttaatggc tcaaaatttc tcattcaaat 56760  
 acccttttgt agatggtcaa ggtaactggg gttctcagga tgatccaaaa tcttttgctg 56820  
 caatgaggta tacagaatct aagttaacta aatttgcaa tcttttaata tctgaattga 56880  
 agtctggaac agtcgattgg cagcctaatt ttgatggctc tcttttagag ccagtaattt 56940  
 ttccagccaa actcccatct attttattga atggcacttc tggaattgct gtaggaatgg 57000  
 caacagatat tccatctcat aatattaatg aaattattga tgccacagta catcttattg 57060  
 ataatccaaa atcacagttg gttgatttac tcaagataat taatggctct gatttctcaa 57120  
 ataattcgcc aataattgct agcaaagatg agctgaatga aatttattcg actggaaaaag 57180  
 gcggcttcaa agctcaagcc caatgggcgc aggataagaa tcaaatcatt atcaacgcatt 57240

09/847,513, 00000001

taccttatca agcatctggg tctaaaaattt tagagcaaat agctgatcaa atgcttaaga 57300  
 aaaaaattcc aatggtggtt gatcttactg atgaaggaga ccacaaggag ccagtaaggc 57360  
 ttgtcataac tttaaaatcg aacagagtaa atgctgaaga tgtaatgaat cacctttttg 57420  
 catcaactga ttacaaaaa aattatagag taaatatgaa ttgattttca ttgaaagggtg 57480  
 gaccaaaagt tttctcccta gttgatttat tgaagaagt gctagtcttt agaaaaagaaa 57540  
 ctgtaataag aaaactagaa catagactcg accaggtaaa cgataggctg catatccttg 57600  
 aggggttatt aattgtttat ttagatttag ataaagtaat aaagattatt agagaatcag 57660  
 atgagccaaa aaaagatatatt attacagctt ttaaaactttc tgatatccag gcaaatgcaa 57720  
 tccttgagat taggttaagg caactagcca agctagaaca aatcaagtta gaacaagaaa 57780  
 gagatacttt ggtagcagaa caaacagata ttgaaaaaat actaagctct aagacaagac 57840  
 ttaaaacatt aattaaaaat gagcttattg aaataaaaaga tgagtttggga gaagtaagam 57900  
 aatctccaat aagagaagca acagaagcca aagtttttttc tgaagaagaa actcttgtca 57960  
 ctgagcctat tacagtagtc ttgtctgcag ctggctggat cagaagtgc aaaggccatg 58020  
 agatagacc cagctcactt tcctatagag gagaagatgt acttcaagat tatggaagag 58080  
 gaaagagcaa tcaagtttca gtttttcttg actcaaatgg gaaggcttac tcacttgcaa 58140  
 gtcactctct tccatctgct agaggaatgg gtgatcctat tacaggaagg gtatctgcag 58200  
 attctggagt aaagttcatt tcttcattga ttgggaatga tgaagataaa ttcattgatta 58260  
 tgaatactgc tggatatggc tatatttcag agtttaaaaa tatgggttct aataagaaat 58320

09/847,513 DeLong et al.

caggaaaaagc atttatgaaa atccctcatg aagcagacct tcttaaaagcc attaaagtaa 58380  
gagacgatca tttgtatata gcagcagttt caaatattgg cagacttttta atttttaaga 58440  
ttgatgaatt gccaactcctt ggaaggggca aaggggaataa aataataaat ataccaaccg 58500  
ctaaatttat agcaaaaagaa gagttaatga cccatgcgca acttgtttct gaggctagct 58560  
ctttaaggat tgaagcgga aagagattcc tcactttaaa actcaaaagat ctagaaaaact 58620  
atatttctac aagagcaaaa aggggaaata tgcttccaca aggatatagg aaggtagata 58680  
aaatgattga agaggttgag ttagaagtta aagaagactg attatagatt tctcaaaaaat 58740  
ttttaatccc tcatcaacta gctcattttc tattattagg ctgggagaaa acctaacagt 58800  
agatgcattc gccttaagaa tcattaaccc attatgatgg gattttttta ttaaatcattc 58860  
tatttgaatt ttgctatcct tgcttacttc cactccgacc cataaaccag cagaagtatt 58920  
tttttcaaaa catttgtgct tctcatttat tttattcaac aaattgataa atctaacttc 58980  
tttctttaaa accttattta aaaatgattt tttagaaatg gtatctataa cttcattacc 59040  
tatggcaat gcaattgggc ctctccaaa agtcgtgcca tgagaccctg ccgacatatg 59100  
ttttgaaact ttatctgaag taagaatgcc tcctagtggg aagccattag atataccttt 59160  
agcaaaaaca agaataatcag gagtaatggt aaattgctca taagcaaaaga gtgttccagt 59220  
cctgccaatc cggattgaa cttcatcaat tatgacaagt gccttatggt tttttgcaag 59280  
tttctttatt ttgtctataa attttttatac tgctttttgtt attccagatt gccattgaac 59340

09847.513 DeLong et al.



cagtttctaataaaccgctg cagtccttatac tgaaaaaacc ttttctagat tagtaatgtc 59400  
attatatgga tgatttttta tccctctagg cagaggcgca aaaccgtctg taagatgctt 59460  
agcttttagct aaagctatac caagcatagt tcttccatga aatgacgttg aaaatgaaat 59520  
gacctcattt ttatttttat taactgttga tgagcaaaat tttcttgcta tcttaactgc 59580  
agcctcaatt gattctgccc ctgaattgca aaaaaatact ttatctgcaa aagaattctt 59640  
gcataacttt cttgcccagg taactgaggg ctcatattatg tagagatttg ataaatgcca 59700  
aagctcctca gattgctttt taagaatttt tattagatct ttatttgaat gaccaagggtt 59760  
agtgacggca atgcctgctg tgaaatcaat atatttctta ttatttaggt ccatacatg 59820  
cgatccactc gcttttttta caacaaaatc tgcaggagcg taaaaaggca ccatataatt 59880  
agtatatctt ttaattatct ctttcataaa tatattgtat cgtggtttca gttcaaaaat 59940  
taaatgacca aaaatttcta gtatctttaa aaccaaacag ctcactcatt ggctttaata 60000  
gaataatatt tataatctagt atatctatg tatgcgggtg catagcattg atatttttct 60060  
tttttggagc aacgctcatt ttgccttttg ctggtttaga gcttggtatt ctattcactg 60120  
cattctattt aagtttttaa tggagtata aaaaagaaaa aatatttatt tctcaagatc 60180  
ttgtaactat agaaaaaggc tctaattatg ctgaatataa atgggaagag ttcagggtcat 60240  
ttacctcttt ccaggtttca aaagacagaa gagatcttct taagctaagc tttagggtcta 60300  
aaggcgagga tgttgaggtt ggcagctttc taaatgaaga tgataaaaaat gtattaatag 60360  
aagagctaac tcagatcata gatacattaa atcagattc cttctcaaaag ccagagcttt 60420

U347512.000001

aatttctttt taacttcttt tagctttaat tcttttatatt cagggatgcc atttctaaaa 60480  
tttgaaaatg attctccctg tattaatggc ataagatatg caattgcttt tggggtaacg 60540  
tccattccat ttttagctat aaaagatttt ggtaattttt tttctaaatt agctattttt 60600  
gccagtggag ctggttcaat ttccaacctt tatttttttag ctttacctct aactattatg 60660  
ggcataaac cattcatgcc ttctactgca tactgaactg ccttggcccc aacagccata 60720  
gcttgctcta ggtctgtttt ggaggcaata tgcgtgacac ttctttgttag atagtcagaa 60780  
acagcccaat gatttttttag ttttaattta tcagtaatta agttggcaat gtaaggcgca 60840  
actccaccaa gttgagcatg accaaatgca tctattgttg ctgattcaga aagaaatctg 60900  
ttcttattat ttttcaacc ttcggatacc acaacaacac agtagccatt ttttttaaca 60960  
acacttttaa cttctgctag aaattttttt tgattgaagg ttatttcagg taataaaatt 61020  
atatgagggtg catctccttt ttcttttctt gcaagagaag atgcagccgc catccagcca 61080  
gcatgcctgc ccataacttc taatataaaa acttttgttg atgttgctgc cattgaccga 61140  
acatctaatag aggttcttag agttgatgtt gctatatatt tagccgcga cccaaatcca 61200  
ggacaacagt ctgttaatac caaatcgta tcaacagttt ttgggatggc tatacagtta 61260  
attggataat ttaatttttt acttatctgc gaaacttta atgcagtatc agctgaatca 61320  
ttaccgccat tgtaaaaaaa atatcctata ttgtgcgctt taaaaacatc aataagcctt 61380  
ttgtactctc ttctgcttga ctcaagatct ttttaactta acctacatga cccaaaagcc 61440

6047513\_000001

ccaccaggcc tatatttcaa tgactctaata gcagatatg attcttttga tgtatctatt 61500  
agttcctcat tcagcgcgcc caatatcca ttcttgccg cataaaatttt accgatgtct 61560  
ttatgcttct tggcctctaa aatcagagcg ccagctgtag cattataaac tgcggtaacg 61620  
ccgccggatt gagcataaaa cgcatttttt ttcatTTTTT ctccatcaaa tgtactgaaa 61680  
taaaaaatat taaagtatta taactaatat gcgaatacat atcttaggga tctgtggrac 61740  
ctttatgggc ggccttgcta agatacttaa agagtcagg catgaaatat ctggatcaga 61800  
tattcaattt tatctccca tgtcagacta tcttgatagc ttgatattg aaatgatcaa 61860  
gggctatgac ataaaaagca tgccagatgc tgacttgtat gtgattggta atgcactttc 61920  
tagaggaaat gaatccgttg agcatatttt gtccaattct cttccgttta agtcaggccc 61980  
tgaaatgctt ggagaaattt taaaaataa aaaggtttta gctatatctg gaacacatgg 62040  
taaaacaacg acttcatata tgttgactca tattatgta gtcagggta gggatgtcgg 62100  
ttttctagtt ggtggaatat caaataatat atctggctct gcatgtcttg ggtcagatgg 62160  
aacttttgtg attgaggcag atgaatatga ttctgcattt ttgataaaa gatcaaaaatt 62220  
tattcactat tcaccaagca ctatagtcac caataatata gaatttgatc atgctgatat 62280  
ttttaataat ttagatgaca ttaaaagaca attccatcat ctaattaaaa taattccatc 62340  
aaatggaaat gttgttttatt ttgctgatga taaaaatata agagatctta tcgatatggg 62400  
aatttgggtca aatcagatag caatcaataa taatgctcat tcaattgagg cagttttattc 62460  
tgataagact ctaaaatatg aagaaagtat ttattcata aatgagttac ctttaaatagg 62520

09/847,513 DeLong et al.

agaacataat tttaaaaaact acattttcggc tatttttggcg gcaaagacag atggaattca 62580  
 aattcaggat tctatcaaat cattagctag ctttgatggc gtaaaaaagaa gattagaata 62640  
 taaaggaagt ttatgatggca tagaaattta tgatgatttt gctcatcatc caaccgcaat 62700  
 agaattttct tctaattgccc tcgtaacaca aaatcccatca aaaaaatac ttggtctcat 62760  
 tgagctaggc tcaaacacta tgtctggggg ttctcatggc ttgtcttttg ttgaatctgc 62820  
 aaaatcttta actcatgtta tctgggctaga tcgcaataat gttttgtcag agaattgctag 62880  
 cattgaatct actaacacta ctgaagattt tattttctgca gcgatatctg ctttctcaga 62940  
 ttatgatatt gttattttga tgaccaataa agacagccaa aaaatatata aaccatttgt 63000  
 agatcacctt gaaaaataat aatttaccag tttttccttt aggaatagtc gccctcccag 63060  
 gtagcatcca atctcttcaa attttttgagc ctaggatat acagatgggt aaaacatgtc 63120  
 tatccaagaa ccatggattt gtaattgttt ttaatgccaa taatgagtct caaggcgatt 63180  
 tcactttttc taagaaggga agttttgttg aaattataga ctttaataat ttgccaaatg 63240  
 gccttcttgg gataactgta aaatctataa ataaggatgat aatcagtaat atatgtcaat 63300  
 tagaagatgg actgcatatt gctgatatta aggacagat agatccagag gtagatgac 63360  
 aagctgtttt ggcagaatat cctgagatat ctaggattct ttctcagctt gtaaagcatc 63420  
 ctaagattag tgacctgccc atccagggtg actttggctc tgctgattca gttgcatacc 63480  
 acttagcagg ccttatacct ctaagctcaa atgagaaaca aaaactatta gaagcattcg 63540

62580 62640 62700 62760 62820 62880 62940 63000 63060 63120 63180 63240 63300 63360 63420 63480 63540

atgcagcaca gcggatgaga attcttttcag actatatattga aagaatatctt actacataaa 63600  
 ttattttatt taatattatt ggcggcttat tagatttttg tctaagccaa ttgatggact 63660  
 taaaaaccgt agggattttct acaaacccttt ttccaatata gtatttgcct ggatagttac 63720  
 taagcattag cccaataata atagtgaaca aacctggcca ggtaacacca gcatcatgat 63780  
 tcccccaata ataaggggaat aacctaaaat atttttaatc agtaaaacta aataccacaa 63840  
 caacgggtta ttggtcttaa atttcgagac atctttcttt ataaaatagt cacttggaat 63900  
 taatccagca agccacctca tactaacgag actaaataaa aaaataaata aagatatatga 63960  
 gctcagccaa agtataagaa ctttataaga ttcaaaaaaa attatgatgg catttaggggt 64020  
 attaaattcc ataattaccc acctttttat ttaataatga tttagtatag caatgtcttaa 64080  
 aaataaaatt aaacattaaa taaatatata ataattttgc caaaaaaagc cttaaatgat 64140  
 atttgactg aaaccctcag tgataataaa gccgaaaatg ttctatcact tgatatcaaa 64200  
 ggtattttctt cttttgctga caacatcatt attgcaactg caaatcga taggcattgca 64260  
 aagtctttat ccgaaaaagt agttgaggag ataaaagcta ataaaattag tatcatgggt 64320  
 gtagagggca agacagaatc aggttgata ttagtagact gtggtgaggt ttagtaaat 64380  
 attatgaaga atgacataag agagttttat gatttagaag gtctttgggg tgaaaaacacc 64440  
 ctcatcgatt cttcgaaaata aatgctatta aatatcataa gtgttggaac ctcaccttca 64500  
 tcttgggagt taacaggaat agaattattac acaagcaaa tccctaaaga agtaagtcta 64560  
 aattttatta acgtaaaagg gcagcaacac ccaaaaagat ctacagaaga ggtcttaaaa 64620

09/847,513 DeLong et al.

cttgaatcta agttaatctc gtccaaaata gattctaata gatatattgt ttgctgggac 64680  
 tcaagtggcg agtcgttgaa taattttgaa ttagtaaat ttttgaaaa atccatgctg 64740  
 gaaaatatga agctttactt tgtaattggg ggttcatttg gcatacctcg agatattctt 64800  
 gataaatcta ataagataat ctcgatctca agtctcacc tacctcacag gcttttcaag 64860  
 attgttctta tagagcaaat ctataggta ttttcaattc tcaaaaatct tccctaccac 64920  
 aaatgattga tctgaatgag aggtttgttg aaaaaagaag ctttttcaat agattgttgc 64980  
 ttatatattt tttctttggg gccctgtttt tattcttgct tttcaagacc tattctttac 65040  
 aggttttctag ctatttctgac tatgaactag cagcactaaa gaataaaaaca aaagaagtgc 65100  
 tagttcagcc tgttcgagga gtcattctatg acagggaatgg aaatatctta gttaataatg 65160  
 tccctacata tgaccttaata atccaaccat caaaaattaa aaacttagat gatttcatta 65220  
 ttaatatattc taaaataata gagctatcag attcagaaaat tgaaaacatt attaaaaatt 65280  
 ttaaaaggag tgcaacttat aatagagagc taacaattaa aaaagatctc tcgaaagaag 65340  
 aaatagcaaa atttgaagta agaagctacc aatttcagaa tgcatttata gatgttaggt 65400  
 atagtagaga aaataaatat ccttacttat tttcacacgc actcgggttat gtagggtggtg 65460  
 taagtaatga taagggtttta tcaatttttaa aaatcaaga tttgaagcaa tctgaaacaa 65520  
 ctttttaagta ttcaggttgtt tttatttgctg gaaaaacagg attggaaaat atttatgatt 65580  
 cagctttaag aggttccttt ggaaaaaagt tatttgaagt tgatgcaagg ggaaggcttt 65640

09/847,513 DeLong et al.

taaaaagagct aagtttttgaa aaacctatta atggaaaaag cctttttact catttagatt 65700  
 taaattcaca aaaaaaagcc tttgaacaaa tgaataatag aaggggtgct gttgttgcct 65760  
 tagagcttaa atctggttct atagttaacct acctaagcac tccaagcttt aatgtaaatg 65820  
 gtcttttctaa tggcatgtcg tcggtagagt tcgaaaaact gattaatgat gtagacaagc 65880  
 cattttttga tagggccggt caaggtcgggt actctccagc ttctacaata aagccagcaa 65940  
 tagcatttgtt tgggtattaaa gaaaaaatag tagactggaa ctttacactt aaagatcctg 66000  
 gattttttgt attaccagag gatcagagga tttatagggg atggaagaaa ggaggtcatg 66060  
 gaacaattga tatgaagggt gcaatcatag aaagtctctaa tacttttttc ttttctcttg 66120  
 cttataaaggc tgatatattaat aatttaatac gccatctttc tgagtttggt tttggcagga 66180  
 atgttttgtaa agattgtttt aatccagatt ctggattatt gcctacgcct gaatggaaaa 66240  
 tgaataatct taatttttggc tggtttaaag gagataccgt taatttaggt gttggtcagg 66300  
 gctatatgag tgccactcca gttcaattag catattactc tgcatttctg gcaaaaaaag 66360  
 gaaatcttca agagctatca tttgtcgaga gtgacagcct aagcaatact gcttttataa 66420  
 ataatttaag catagataac tcagactggg atcaaaattca ctcaagcatg attggagtta 66480  
 ttgaagatcc gagaggcact gctaaaagat taaaaccttt aaaatcatat gttgttgctg 66540  
 caaaatctgg aacgggttgag cttgtcagca cacaacaaaa ggaagattac aaactagtaa 66600  
 ggcaaaaatat tggtaacaga gaccatgcaa tcattgtggc ttttgggtcca atgcctgac 66660  
 cagagtatgc agtaagtgtt gttattgaaa atggtgaaa cggcgggttct gttgcaggtc 66720

09847513.0000001

ctgttgctat tgctgtttta aatgagctta tcaataaatg aagaaaaaat tagacttta 66780  
 aaacttttagc atttattttg atcaatatatt atttattgcg ataaccctgt tgtcggtcat 66840  
 gggtttatatt tttttatata gcgcatacaca agaagatatc agcactgttg ctaagcaagc 66900  
 tgtattcggt gggtttgggt tgctgttaaat gtttgtagtt agccaacctg accctgattt 66960  
 ttataatata ttttctgggt tattttttgg ggggggastt gtattgattt ttctaactat 67020  
 gatttttggt aaagaaaataa atggagccaa agatgggctt gatttaggat tttttaccct 67080  
 gcaatcctct gagattatta aaatatcatt gccaattttt ttatcatcat atttatataa 67140  
 taagccactt ccaataagca ctaggcatac ttttattaca ttgatattaa taggtttttat 67200  
 atttgactt gtagctagac aacctgactt gggaaccagt ttagtagtat ttatgtcagg 67260  
 gggttatgta ttatttttag caggattaag ccgccgttta ttggatctgc aatagcctta 67320  
 tttttattat ccttacctbt tttatggaat aacttttctag agccatttca acaacagaga 67380  
 gttttaactt atttgaacc agacgcagac cttatggta ctgcatggaa tataactcaa 67440  
 tcaaaaatag caattggatc aggcggaatt aatggaaaaag gttatcagga gggctcccaa 67500  
 gcccatcttg atdttttacc agagacagaa acagatttta tttttgctgt tattgctgaa 67560  
 gagtttggtt ttattggagt ctgtattttg ttatcagtat ttdtctttat atkactcaga 67620  
 tgtttatatt tagcatttaa tgcaagagat agattttgca ggttaactat aggaggccta 67680  
 agtttagttt ttgcctctac attatttatt aatttagcaa tggttgttgg tgtagttcct 67740

09/847,513, 00000001



gtagttggta tgcctcttcc atttatcagc aaaggcggct catcttttgct atcctkttat 67800  
 atagctttttg ggattataat ttctatggca acacataaaa aattaatgca aagatgaaaa 67860  
 aaattatatt tataacttta atattcacta ttccaattac tgcggattat tcgaatcatg 67920  
 aagatagcca aatggtaata aatgaactcg tcacaaaaaca cggtttttgag gaatccttatg 67980  
 ttactgcaat cttaaaaaat gcaaaaaaagc gtgatgagat gcttaaatctt gttgctaatac 68040  
 ctgctgaaaa aacaaaaaaca tgggatgaat atagggctat ctttataaaa acaaaaaagag 68100  
 tttcagaagg caaaaaatctt ataaaaaaaa atattaatgc tttagagagg gctgagaaaag 68160  
 aatttggagt tcctaaagaa ataattactg ctatttttagg cgtagaaaact aattatggca 68220  
 gcaataaagg tggatacaga gtttttagata gtcttactac tttaggtttt gatgacccgc 68280  
 gaagatctaa ttttttttaga agagaactta tagagtctct ccttttaaca agagaaaaata 68340  
 atttagatat caaaacaaca aaaggctcat atgcggggagc catgggatat gcacaattta 68400  
 tttcgtcaag ctaccgagcc tatgctgtag attttgatga agatgggttac gttgatttat 68460  
 ttaattctgt tgatgatgca attggggagta ttgcaaatata tcttttatgtt catggatgga 68520  
 agagagaagg aaagatcgta acaaaaactt atccaaacaa tgttagaaaa ttttataaac 68580  
 ctcatgagtc tctaacaagg ttcataacctt taatcttttaa tgaagatgga aaagatctttt 68640  
 tttttattgg tgatgataac tttagggcta ttgctaagta taatattagc gatgtctatg 68700  
 caatggcggt ttattactta tcagaagagt taaaaaaaaatg aaaaaattat tattcacatt 68760  
 actatcaact tcaatatatta ttcaagcaca gagtttttgtt ccagattctc ctgaggttaga 68820

09847513.000001

tctgaagagc tatactcctaa ttgagccaaa caccaatact gttattgcag aatttaattc 68880  
agatttgga atagaaccag ctagcatgac taaaattatg actagctatg ttgtggctga 68940  
tcagattgca aatggtttta tatctcttga tgacgaagta ctaattagtg aaaaagcatg 69000  
gagaatgcaa gggctctaaga cctttataga agctggaaaa aaagttacgg tatctgatct 69060  
tttaaaagggt attatgattc agtcaggcaa tgatgcctct gttgcaattg cagaatatgc 69120  
aggcggcact gaaagaggct tcgttgactt aatgaactct tatgctgcct ctttggagat 69180  
gaataatacg atctttcaga attcaacagg gcttccagat gaaaatcatt ttctatcagc 69240  
aaaagattta gccaatctga cagctaacta cathtaataa ttcccagaag aatatgcctt 69300  
atataagcag aagcaattta cttttaacaa tattaagcaa ttaaatagaa ataagctttt 69360  
atggagagac gactcagcag atggtgttaa aactggtcat acagaagctg ctggatactg 69420  
tttagttggc tcagcaaaaa ggggcgggtat gaggttatt acagttgttg caggaagtaa 69480  
gtcagataat gatcgttttt tatcttccca aagattactt gagtatggat ttagattttt 69540  
tacaacacag aagatgttaa gtgcagaaaa agaataatcag aatattacag ttgggggtgg 69600  
acaagaaaaag atacttggcg tgggagttct agaggatata tctattactc ttcctagaac 69660  
aagtttcaa aatgttgaaa ctgttttata agttaataac aatatccagg ctctatcgt 69720  
agttgggtcaa aaagtcggga ctctagaaat aattagtaat gatgagattg ttctagttac 69780  
ggatttggtt gctttaaaaa atattgaagc taaagtttt ttgggaagaa tatggtcaaa 69840

09847513\_000001

gtctgttctt tggatattta gcttatttgg tctaacagat gaaaacacca cttaaaggcg 69900  
 tttttaatgg ggcttttgat actgtagata acataaaaaat atcacctttt tctcgagctt 69960  
 atactttttc tgatagcgta tatgaagtag ttcctttctt taattcaact gcaattgctt 70020  
 ttgatgatca tataaaaaaga ttagaatttt ctgctagtca actagcgatg gccgtagatt 70080  
 taaaaagaat tataattgaa attaatctt ttataacatc atctgaattt agtaattggct 70140  
 atgtctatta tcaagttact cgaggcggtg atccattaag gtctcatatt catgagccaa 70200  
 atttaaaaaat agaaaccttt ggctatgcaa aggcgcactt attccaatgg aaaccattaa 70260  
 gagtatcagt atgtgatgat attaggtggg ggagggtgtga tattaaatca acatcattgc 70320  
 ttggtaatgt tatgaacatg aatgcagcaa aactagataa ttgtgatgaa gttattatgc 70380  
 acaaagataa tttgttgacc gaggtggcg catctaattc ttttttgtta aaaaatgatt 70440  
 caatctgcac ccagctcta aatggaaaata tacttcccg cataaccaga gcattactaa 70500  
 ttaatgaatt aaaaaattat agtatagagg tcatcgaaga taacttcagg ctagaagatt 70560  
 tatcaaatgc ttcttgcgcg tggctcaca gcgcaacgaa agggctggcc cccatttctg 70620  
 aaattagtaa cctagaatct cacttggatt tataccatcc tcttttcaa aagagtgaag 70680  
 aaatttttaa taagaagtcc cttagttaga agtcagtta ataactatat tatcaagctc 70740  
 gtaccaaaaa tctttatcac taagaccttt atttgcaaga tcaagctcaa atacattttt 70800  
 ttggagctgt attaaagttt gtagtttatg atttttaata aaatgcatat attcagaagt 70860  
 tttatttttc caaacccag atttaattag actatctaca ggattttttac tttgtctagc 70920

09/847,513 . 0000001

attggctgca ctgttaataa ttttaccaac tatccaaact agaagaggtg cataatgatc 70980  
 ttccgctgat tttattgaat ggataatttt taaagcttgc ttggtattat ttgaaataat 71040  
 tttatcttca agctcaaatg gcataaattc tgcagactcc acagaagggtt ttctttcttg 71100  
 gccatcacca tctttataag tcagtcttag aagctttact tcattctgtt gggcaactaa 71160  
 gtttccagag ttcatctcag atatatcttg cactaaagag ggcctatctt tttcagaaaat 71220  
 aaaagatagt tgatgcttaa gccaaatctt ttcttcaaaa gatttaagtt ttccacaatc 71280  
 aatgatgagg gctatctcat ccatctgttt taccactta gttgtgaaat taagcttttc 71340  
 attgcttgaa ttaataatta ttgcaatatt atccatatgg ttaatatctt caatttgaga 71400  
 tattttaatt atttgggtcag gtatttttcc ttgatcatga ttaatctcaa ctattatttt 71460  
 tgagccaaat aaagatccag aagcattttc aataatagtc tgatctattt tatcaaaacc 71520  
 atcttttatta ataattgttt ttcttttgaa ccttttattt gataaatgtt tgagcaaaaag 71580  
 atcttttagag ttattttttta agacaacctc agatccaaaa ataaagaaga tattttgtga 71640  
 ctgatctaaa tattttttgag ttgtcattgc ctacatttc aaaagcatgc acctcaaaaa 71700  
 ttatttgatc aattaattcc ttctgcattt cagactcaag agatttaatc atttgatctt 71760  
 gtgcaaatgg attaaagtca ttcgatttat atctcttgta gctaacaagc tttttactta 71820  
 tttgttttatt tccaatcatt tttatatgta tttctaaaatt aatttctcct tcaagagctc 71880  
 ttagagaaga gcctccatag atatcatatt tatttaaaat atagtcttgt attagaatct 71940

gattttgacg attctcagct tcttttttctg tattgaagag gattgctaac ttttgctcaa 72000  
aagaattagg gacactacta tcaaacccgaa aattaaatag atcttttttta tctagatcta 72060  
tatttagttg attaaaattgg cagccagtta aagatatcaa caagcaagtg tataaaaaaga 72120  
aaggctcttag tttcataacc tcatagtagt tcattgaaga atatatatca aagattaatt 72180  
ttaaataaca aaatttatta ttttttcttt aatataaata gtttttttaa tagccgagtt 72240  
gtttattgtt gatgccacat tgtcgatagc cagtgctaaa gcttctatat ctttttgctc 72300  
aagggttttg totatcatta ccttacctct caccctttcca tttacttgaa ctattaaatt 72360  
aaattcagaa acttctaata actcctctct aagaacaggc caagaagatt caatttcttc 72420  
ttgtgcaaag tcaaaagtaaa aattattcca taaatgttgt gagatatgag gcgcaatagg 72480  
attgagagtt ttaagaataa ttattattgc ttcattaaga caatattgat ttgagatcga 72540  
cgcattacct tctttaaaag aatcagggtat aaagttaact agctccatta tcgaagcaat 72600  
ggctgtatta aatgaatatc ttgtttcaaa gtcatttgta acctttttaa gagtattatg 72660  
cgattttctt cggagctcta cttcttggtt cgaaggatcc ttgggctcct ctaagtcaat 72720  
aaattttcta ttgctgacta agttccatat ttttttcata aatcttgagg caccttctac 72780  
agatgactct gaccattcta gactttgctc agggggagct gtaaacatca tgtaaaagcct 72840  
tacagtgtca gcgcatact tttctatata agattgagga tctacagtat ttccttttga 72900  
tttgacatt ttagcccat ctttaagaac catgccttgt gtaagaagct ttttaaaagg 72960  
ctcattgcct tcaactaaat ccatatctct aagcgcctta tgaaaaaatc ttgaatatag 73020

09/847,513 DeLong et al.

taagtgcagt atcgcatgct caatacctcc tatataataa tcaacaggca gccaatatatt 73080  
 tgaattttta tcaaacatttt catctgcatt gtcagatgaa gtaaatcttg cgtaataacca 73140  
 tgatgagtcc ataaatgtat caaaagtatc tgtctctctt tttagtcgat cagaaatatt 73200  
 ataaaaatct tcatttttgac ttaagggaat aggtgccgag ttcttttttta gctctggaag 73260  
 cactatagc atatcctttt catcaataac tcttggttca ccatttttcat aaacaaccgg 73320  
 aatagggcaa ccccaatatc ttgtctgct cacaccccag tctcttaatc tgaactggat 73380  
 taactgctca ccaagggttg cgtctgctaa atctttaata atttcaagag atgcctcatc 73440  
 agagtccatg ccatacatatt tatcagagtt tattaatttg cttttttgta caactggaag 73500  
 ttcattatta ttatcaaaac taataacttg aggaatctct agattatatt tagatgcaaa 73560  
 ttcgaaaatct ctttgatcat gagcaggaac acccatcact actccagttc cataatcaag 73620  
 aagtacaaaa tttccaatcc atacgggtat cttcttcttt gtaattggat gaattacatg 73680  
 cattccacta ttaattccaa gttttttcagc tttagccata tcagcttcag cagctttcac 73740  
 ttctttgcat ttgtttagaa agtctttttat cgattcatta ttttttgaca accctattga 73800  
 tatagagtga tttggtgata ttgctaaaaa agaaacacca aaaattgtat caggctcgtg 73860  
 tgaataatacc tttaaagaat catcggaatc ctcaatcatg tattttaattt ctgctccctg 73920  
 agatttccca atccaattcc tttgcattgt ttttaacattt tcaggccaat ccacctcatc 73980  
 caaagatgtt aaaagttcct cagcgtagtc tgtaattttt ataaaccatt gatcaatttc 74040

09847513, 0000001

ttttatttca	acttgcgccc	cagacctcca	tcctttttcca	tctataactt	gctcatttgc	74100
taaaactgtc	tcatcaactg	gatcccagtt	cactaatgat	tttttcctat	aaaccaaacc	74160
tttatcataa	aacttttttaa	atattaaactg	ctcccactta	tagtattccg	gttcacatgt	74220
tctaagctct	tttgaccaat	catacccaag	gcccaaaagt	ataagttgct	gtttcatatg	74280
ttcaatattt	tgattcgtcc	aatcttttgg	gctaaacttta	ttagcaatcg	ctgcgttttcc	74340
agcaggaagc	ccaaatgcat	ccagcccat	aggctgaaaa	acattaaagt	tattcattct	74400
tttatatcta	gatatgacat	caccgatcgt	atagtttctt	acatggccca	tatgtaattt	74460
acctgatgga	tatggaaaaca	ttgataggca	ataaaaattt	tctctaccat	caggattagc	74520
tttaaaattg	tcttctttta	accattcatc	ttgaatggtt	ttctctat	ttt gacttggggtt	74580
atattcagga	ttcatTTTTT	taggaactca	ttttctaaagt	aattaatagt	atgtttat	74640
tctttaaagg	tatcatcatg	cagaatcctt	tgatgaagag	ggtgatttgt	tttaaatgccc	74700
tctataaaaa	attcatctaa	agcactcagc	attctcttaa	tagccgagct	tctagaat	74760
gcctgagtaa	tgatttttgc	tagaagtga	tcatagttag	ggggaactct	atatacctcca	74820
taaatatgtg	agtcatatct	tatgccaaaa	ccgccagggtg	tatgcatttt	tgtaattggtt	74880
cctgggggatg	gttgaaaaatt	atcaggatct	tctgcattaa	ttctgcattc	tagtgagtgg	74940
ccatgaaaat	taatatcatc	ttgatttagc	tcaattggca	tttcaagagc	aattcttagt	75000
tgtgcctttta	ctaaatcaaa	gccagttatc	atttctgtta	ccggatgttc	aacctgaatt	75060
cttgtattca	tttctataaa	atagaattga	ttatcttcgt	ataaaaaattc	aatagtgcca	75120

U3042513.1000001

accccttcgt aatttatttc ttacataaaa ttgacacaaag ctattagtgt tttattaaga 75180  
 gcttcttgat ttatattaag tgctggagct tcctcaatta ttttttgatg tctcctctgc 75240  
 atactgcaat cccttggtacc taaatgtatc gcctttcctt taccatcacc aactatttga 75300  
 acttcaatat gtctagggtt cccaataaat ttttctagat aaatggtttc attgccgaat 75360  
 ccatttttag ctccctgcat tgttatttct gcatgtccaa taagatcttc ctctctttca 75420  
 acaactctca ttccctcttc accgccacca gccgttgctt taatcataat tggatagccg 75480  
 atatctctag caattttttt aaactcatca ttatcagatg gaatctcatc ttataaacct 75540  
 ggaacgattt gaattccaga tttttcagcc agtggttttg ctgttatttt gtcgccatt 75600  
 ttttggattg tttctgaggt tggcccaata aatttaaata cacttttttc acacatttct 75660  
 gcaaaattat gatcttctgc aaggaaccca tatccaggat atatagcatt agcacttgta 75720  
 agttctgcag ctgagagaat ggcaggata tttaggtagc tttgagttgg agatgcaggc 75780  
 ccaatacaaa cagtttcac agagaatctt aaatgtttta gatctttatc agcttctgag 75840  
 taaacagata ctgtcttgat accaagctct ttacaagctc ttatagctcg aagagcaatc 75900  
 tcaccacgat tggcaataag aactttataa ctcattttaa ttgacagtta taattttttg 75960  
 tccaaatca accggacttc catcttcac gtcaataactt gaaatcttgc catcaaatc 76020  
 tgatttaatt tcattcatca ttttcatggc ttcaactata cagagcacat cacctacttt 76080  
 tatattgtct cctactttta caaatggatc tttttcggga cttgggcttc tataaaaagt 76140

09847513.080801



tcctactatt ggagaggtaa caacatcccc tatgacagtt tgcttagcct catccaattg 76200  
 atgagttggt ggtgagatag ctggctcatt ttgacaatg attggttttg aaacaaaattc 76260  
 ttgattagag ccgttatctc ttgaaaattct tactgactca tctccttgac ttactttctat 76320  
 ttctttttaa tcagactctt gaagcatttc tataagtttt tttattttcc taatatccat 76380  
 ttattttacc ctgtatttta atatgccctt aagcctgcgc tcttcataac ccttagcacc 76440  
 aataccagtt atcactttct cagctatatc agatagatat gagacatgtc taaattcctc 76500  
 tcttgtatat atgtctgaaa ggtgaacctc atagaatggt atgttgacac caagaaaatgc 76560  
 atcccttata gctatgctcg tatgagtata tgcagcagga tttattatta caaagtctat 76620  
 tttttcttca atagtgtctat gtattgattc tattatttca tgcctccgat tacttttga 76680  
 agatattaaa ttgcattgat tttgatttgc taagactaat aaatcttggt gaatatcttc 76740  
 aagtgatttg ctgccatata cttccttttc tctagtctcg agaagattga ggttggggccc 76800  
 atttataagc aatatattca tataatagat tttaacagaa atttacagat tttttaggat 76860  
 tttaaagata tttttataag agttgatttt ttgcatggg atagcaaaaa cccgcttcag 76920  
 agcatccctg atagtatat attaacctccg ataagttcaa tgaagaattt atttcaatgc 76980  
 ttataataaa ttcatctcta agaattctcag tctcgccaaa aaattcatcc ctatatagat 77040  
 ctttacttga ctcaagcgtc ttaaatggga gcactttatc tttaaatttg agctctatac 77100  
 tatccaagta catataataa ccattctctta ttbgccaagt tatatatgct gcttgttttt 77160  
 caacattggt tgttaaagca aatacttcat tggcttctgg gacctattt gacttttcaa 77220

09/847,513 DeLong et al.

aaagatttgc tgaattcaat tcacctgcaa tcacattgga tagtagtata attattaaac 77280  
ttttgatcat gaagacttaa gaataaaacc ctaagtactt aaaataaaga ttttattaaa 77340  
taacacggta gttaaaaaatt tatgttccag actttcgac taaaaatttt cttttatatg 77400  
cctatttggc ttttaaaaat aattttttat agaaaaagaa cagtaataag agggcatcaa 77460  
tttgatgctc aatctgctgc attattgtct ttgcttccaa agaaagattt atctgaatta 77520  
ttagatggtg aaattgcaaa agctagaatt actcttgaag aagcaagaat tcaaaaataaaa 77580  
gtttctttta caccgtctat acaagtcaga aaagtagatc atattttacc aaagcatgat 77640  
ttaattctca gagaatacaa gccgcatcaa gaggatttaa aaaaagttat tctctatttt 77700  
catggaggag gttatgtcct taactcagtc aatacacatg atgatatggt ttcatatatg 77760  
tcagataaaa taggagtaaa gttttattct ctagactata gactatcgcc tgagagcaaaa 77820  
tatactgatt ctcttgatga tgctctagat gctttttctt ggctcatcgg ccagggatat 77880  
ggaccaggtg atatttcagt ttgtggagac agcgcaggcg ctcatctagc tgcttctttg 77940  
tgccattacc tagctgaaca aaataaaagat ttgccctagta gccagctact aatatatcca 78000  
atgtgtgato cgctcgtgtc atctgagtct tataaattat tatcttcagg atatcttcta 78060  
actaaaaaaa ctatgatttg gttttgggat aagcttagaa ctcccgaagt aaacaataac 78120  
gattcggcct ttaatctctt aaaattttaat tttgaaaaaa ctttaccgcc aacaattatc 78180  
gttacaggtg gctttgatcc tctatgcgat gatggagaaa aatatgcata tttattacat 78240

09847513.000001

aaaggtaaac ataatgtgaa acaattacat tatccaacaa tgtttcatgg gtttgcata 78300  
 atgactaaat tgaagcagc gcagatagcc gttgaagatt ttttaaaaga atataagaaa 78360  
 atactatgag taaaatttta gaagtaagt acttaagcat caatttttca acaagggatg 78420  
 gattgtttta tgcagtggat aatataagtt ttgatataga aaaaaatcaa accttggcct 78480  
 tggttggtga gtcaggttct ggtaagtcgg taactgctat gtcaattctt cagtccttc 78540  
 aaaaaccaca agcatcatat tccaaagagt cttctattaa gtttaaatggc gatgagataa 78600  
 taaatgcaa gtatgaaaag ttactttcct tgagaggaaa tattatatct atgatatctt 78660  
 aagagccgat gacctcacta aaccttatac acagagtagg taatcagata actgaatcaa 78720  
 tactacttca ctcaaaaaag tcaaaaaaag atgcaataga tgaagcaaaa aaattaatgg 78780  
 cacttggtga gattgatgat gttgaaagac ggttctatgc atacctcat gagctttctg 78840  
 gagggcagcg acaaagagtt atgattgcta tggcccttgt taataaacct gagctattga 78900  
 ttgctgacga gccaaacaacc gctcttgatg taactatcca agcccagata ttagatctca 78960  
 tgtctaagct taaaaatgaa ctaggcatgt caatactttt tattactcat gatctaggcc 79020  
 tcgttcaaga attttctgac aatgttttgt ttatgaagaa tggcaagata gttgaacaaag 79080  
 gaaatactgt tgaggtatgt aacaaccct ctcacgaata tacaaaaaaa ctttttagatg 79140  
 cagaacctca gccaaaatta gataaccctg taagtgatga gccaaataatt gagatcaatg 79200  
 acttaaatgt ttactactca ataccttcaa ctaattttt taaaaagaat acttttcatg 79260  
 ctgttaaaaa tacttccttt agtattata aaaatactac aattggcctg gttggggaat 79320

09/847,513 DeLong et al.

ctggatctgg aaagtcaact ttgggtaagg ctatagcaaa cttagtttct tataagggga 79380  
 atattaaatt tgaggggaaga gatatcaact caaatcccc aaaagaaaaat aaagaattga 79440  
 aaaaaaatgt ccaaatgttt tttcaagatc cttatgggtc attatcacca cgaatgacag 79500  
 tgggggagat agttggtgaa ggttttaggtg ttcactttaa gcttacaaaa aaagaaaagag 79560  
 acgaaaggat agataagggtt ctgtcagatg tcggtatcga aatagtagct aagaataaat 79620  
 atccgcatga gtttttctgga ggccaaaaggc agagaattgc aatgctaga tctttaataa 79680  
 tgaatcctgc ttttatgatt cttgatgagc caacatcagc attagatagg tctattcaaa 79740  
 ttcaggtaat cgatttattg aaagagatac aaaatgaata tgggcttact tatcttttta 79800  
 taagtcatga tttaaagggtt attagatcga tgtcagactt tatttttgtt atgaaaaaatg 79860  
 gagaaatcgt agagtcagga ctttctcaca aggtctttga aagcccagag caagactata 79920  
 ctaaaaaatt actatcagct gctttaaaagt atgcatctga ataattaaat atatymcm 79980  
 tggcaaatag aaartattca aaagagctcg ttgacgggtc taatcaagct gcttctagat 80040  
 caatgcttag aggagtaggt ttcacatctg aggmthttcac aaaaccattt gttgggattg 80100  
 cttccacagg agcaaaaagta acccatgca atatgcacat aaatgcactt tcagagatcg 80160  
 ttgagaaaatc agttgatagt tcaggaggaa aggggtgtct ttttaatact attactgttt 80220  
 ccgatggaat ttctatgggt acacagggtg tgaaatatct tcttgtttct cgagaggtaa 80280  
 ttgcagattc aatagaaact gttgtgggat gtcttggtta tgatggagtt attgctgtcg 80340

09/847,513 DeLong et al. Corrected Sequence Listing (August 4<sup>th</sup>, 2001)

gtggttgatg taaaaatatg cctggatgca ttattggaat ggcaagatta aatagaccat 80400  
 caatatttat atatggtggt tctatcaaac ctagtaaaga aaataccgac tatgttactg 80460  
 tttgtgagaa aactggagag tactcaaaa gcgatcttaa agaattctgaa ttaattcatg 80520  
 tagaaaaaat ttccgtaaaa gggccctggat cttgtggggg aatgtatacg gcaaatacta 80580  
 tggcttctgc gattgaagct ctaggcatga gtcttcctgg aagcagcagt caagatgcaa 80640  
 tttcacacga caaagaagat gattgtttta aggctggcga agcgataatg aatttattag 80700  
 aaaaagatat taagccttca gatattatga ctaagaatgc ttttgagaat gctataacaa 80760  
 tggtaatgc tctaggaggt tcaactaatg cagttctgca tttattggcc atggcgcatt 80820  
 caataggggt tgatttagag ctagatgact ttacaagaat aggaaaaaaa acacccgtta 80880  
 tggcagatct taagcccttt ggttctcatt atatgtctga actcaatgct aatggcggta 80940  
 ttcagccact aatgaaaaact ttgcttgaga agggattact acacggcaat tgtcttaccg 81000  
 ttaccggtca gacgcttgct gaaaatcttt ctggaataaa accttacgag cctgatcaag 81060  
 agataaattaa atcatttgat aatccaatta aatcaaatag ccatcttaga attctgtatg 81120  
 gcaacttagc gaaagatggt gcagttgcaa aaattacggg taaagaagga acttcctttg 81180  
 aaggaagtgc tcgtgtatgtt gattcagaag aagaaggggt taaagcaatc ctatctaaat 81240  
 ctataaaagc tggagatggt gttgtaatta gatatgaagg gccaaaagga ggtccaggca 81300  
 tgagagaaat gctaaaacca acatctgcc aataagggtca aggtcttggc gatcaggtag 81360  
 cttttataac agacggtcgt ttttcaggag gcactcatgg atttgttgtt gggcacatta 81420

09/847,513 DeLong et al.

ctccagaggc tgcagatggc ggcttaattg caataattaa agatggcgac tctatatataa 81480  
 ttgatgcaga ttctgataag ttaattcttta atattttctga ggatgaaatt tcaaatagac 81540  
 taagtagatg ggtaaaccaca aaaacgcctc ccaaaaaagg agtcttagca aaatttgcaa 81600  
 aaagtgttaa atcagctagt cttggagcgg taacagatta aatatgtatt ttaaaagaaa 81660  
 atttccaaat agtagactaa gaagaatgcg tctcaattca aacctcagag acttgcttgc 81720  
 tgaagttagg ttgtctacaa atgatctaata tcagccatta ttataaaaag aaggcttgag 81780  
 tggaaaaagag gctattgaaa gcatgccccaa tattaataga tatggacaag attcaatttt 81840  
 ttcagaaata gaagagctac tagagcatga tataaatatc attgccctat ttccagttat 81900  
 taatgattcc aagaaaaaca gactggaga tgaggctatc aatgcattca atttgatgtg 81960  
 tgaactatc agtaacataa agaaacgatt ccctgaaata atcttaatat ctgatgttgc 82020  
 tctagaccca tataccgac atgggcatga cggtatttta aaaaatgatt atgttgataa 82080  
 tgatgagact ttagctgttt taaggaagca atcactaaca ttagctcagg ctgggacaga 82140  
 cataattgca ccatacagata tgatggatgg aagaataggc tctataaagag aggcattgga 82200  
 cgaaatggat tataaaaaata caatcctact atcatatgca gcaaaatata attcaaagtt 82260  
 ttatggtcca tttagagatg ctgtttaatc agcttcaagt ctaggcaagt cttcaaaatc 82320  
 cagttatcaa atgtcaccaa aaaatatataa tgaagctcta catgaagtgt ctatggatat 82380  
 taatgaaggt gcagacatag tcatgggttaa gccagggtatg ctttatctag atataatttc 82440

09/847,513 DeLong et al.

aaaagtaaaa gaaaccttca aagtacctac ctttgcatac caggtagtg gtgaatatag 82500  
 catgcttaaa ctggcgattg ataaaggatg gcttgaaagc gatgttatgt tagaatcatt 82560  
 aataagtttt aaaagagcgg gagctgatgc aattctaacy tacgcagcta aagaaatttc 82620  
 caaggagata actaacaaat gagcaatgtg atagaaaattc gtgatgaaga aagctttaat 82680  
 agcgacgtct taaattcaga aaaacctgta ttggttgatt tttgggctga gtggtgtgga 82740  
 ccttgtaaac agcttgcacc aactgttgaa acagttgcag cagaaaaatc agaaacatta 82800  
 aaggtttgca aaatggatgt tgattcaaat agagagattg ctgctaaata tggaataagg 82860  
 tcaatccctt cattaattat atttaaaaac ggagagcctg caggagttga agtaggtgct 82920  
 ctaaccaaac acaattaga ggactttata agtacagtag tttaactttg caaagacttc 82980  
 tttgcataat gcaagaaaata ggattatcat ttcgacttct aaggctgaaa gccactcaaa 83040  
 acaaaaacc accttttcaa tcataaaaac acaactaga acggaataa ctaaatgaac 83100  
 cttactgaaa ttaaaataaa accaataaat gaacttgtag atatagctac tgagcttggg 83160  
 cttgaggatg ttggaaggct gaaaaagcaa gagataatat ttagaatat taagcataag 83220  
 gcttctgaag gtgttgatat ctatggtggc ggagttcttg agattttaaa tgatgggttt 83280  
 gggtttttgc gatccccaga aggtcttat tgcgctggcg aagatgatat ctatgtttca 83340  
 ccaagccaaa taagaaagtt tagcctcagg aaggagact cagttgctgg gaagataagg 83400  
 acccctaaag ataaagagcg ctattttgca ttaatccaag ttgatactat taatggtgaa 83460  
 gagccaagaa agactaaaaa caagattctt ttgaaaaatt taactctct ttttcccaat 83520

09847513 . 0000001





tactaaaatt atagattaaa aaaatttcat tattatctgt aacttttggt atttgctcca	84600
gcaaatagat tacttcatag caaactatct catctagact agccttaagt ttttggtgaa	84660
gaaaaccatt tgaatttttcg ccgcaaaaata aaagactctt tcctagaaaa tttttctctca	84720
ttaacttaag aattccttca gatgaatggc tttgtggaaa atgagatttt atgccacttt	84780
caagaagttt attggaagt gcgggcccaa cagacaaaaa attgtgtgga agatcatcta	84840
aatcgaaaaa agatttttaa atatcaagtc cgtacgaagc agcagcttgg ctagtaaaaa	84900
ttaagtttga atatgaatga atatttttaa ttttatcgat agcgattttt gaagggtcga	84960
ttgaattaat tttagaaaaga taaatatttt taagagctat ttcttgagct tcacataaag	85020
atattaggtt tcctgataaa tttttgggcc tcgtattaat tatcattttt taaaatttaa	85080
gatttcgctc ctccaaaaat aaactcatca gcaaataaca taatatctat tgcaatatct	85140
tgaaggaag atatcttttc ttataaaatt ttttcgttgc cttcataaga taatattttc	85200
cctgatattt ttatttggcc atttttatct tcacataaaa tagctattgg agacaggcaa	85260
gtaccttcca tggcagcaac aaaagatcgc tctgcaactgg caagaattaa ttcttttaggg	85320
tcaccaatat ttctaaaag ctcaataata tcttttttat ctgacagaca ttcaattgct	85380
atatatccct gagatgcaga aggtaacatt tcttcaattg agaattcata tgaattttgt	85440
tttaaaacca gcctttttat agcagcctta gcaactacca gcccatcaaa cagaccatca	85500
ttgagtttct ttattctagt agctatgttg cctcttactg gaatagtttt tatatcaggt	85560
cttaagttat taatttgtgc ttttcttctt ggaccagaag taccaatagt tgaattttcta	85620

0947513 : 000001

gcaaatcag aaatgattt cccatcctta gatagaagca ggtcttctgg agattcccta 85680  
ctaaaaacac ttattatttt aaacttttga tcaagctttg ctggaacatc ttttaaaacta 85740  
tgcactgcaa tatctgcttc atcagcctct agagaggact caagtgtaga aatgaatagg 85800  
ccctttccac caatttcacg aagaggttta tctgttttgg ctccttctga tgtcatagga 85860  
acaagctcaa ccttaatat attgatctta gctaataact catctgcaac aaattttgcc 85920  
tgatacattg ctagttctga ttgccttgta gctattctta ttttcatttt gtttccaata 85980  
aaagagcctt cacatcgcca acactctttt ctttcaagat agtgaattcc tcaggaattt 86040  
ccaaggtagt aaatttacta tactcaaggt agatttttaca agatgggcttt atttcattct 86100  
ttcgaattat agatttttga acttttaatt catattcttc gccaaaaagga ggatcaagaa 86160  
gaattaaaac aaaacgagat aagtcatttt ttttaatacca actaaatgca tccttaaaaa 86220  
aaacttttaga tttatccttg atgccaagaa gttgaatat ttttgctaac acagaatagt 86280  
ttttttttatt aagttcaacg aacacaaact tttttgactg tctggatat gttcaatac 86340  
caagcgacc cgttccggca acaaaatcaa gacatatata attttcaatc tcaaattgaa 86400  
gccaaattaa aattatttct ttttaatttat ttgaagttgg tcttaaggaa tccttaaaatt 86460  
cgaaaggtat ttttttacct tttaaataac ctccagtgt tctgatattg tttttcattt 86520  
tttttgcaaa atgggtatgc ctatttatta aataaatata attamaaata catttaaaact 86580  
aagttataat tcatccatgm gtccacaag ataatttttag aaaagcaatg aggagctata 86640

09/847,513 DeLong et al.

ttactctgt cagtgatg tcaaatgtg acgagaataa aaagttagt gctataacag 86700  
 ttcttcagt tacctctgtt totttagatc ctctagttt gctcgtttgt atcaataaat 86760  
 cagctggaat tcacaactca ataaaagaag ggctctcttt ttgcataaac cttttaaata 86820  
 aaaatcaaga agatatctt aatctatgca gtctatttaa gaccgaaggt gatagattta 86880  
 atagtgtga ctgggattta agcggcactc cgtttttaaa aagtgccag gctaataatt 86940  
 ttgtactgt tgatcaatta atttcatacc acacacatac tattgtgatt ggtcatgtca 87000  
 caaactctct tagcgatgag aaaattaata cactgacata tgttgatggt agctatgaat 87060  
 aaattttcaa aaaatgtatt ttttatttta ataatactca actctttttt tcttgcctct 87120  
 aacatttttg ctagccagga agagtgtgaa gaaaagccaa gtgttttttat tatctctcct 87180  
 caagacggtt ttatctcaga atctaataat gtaaaagtct tatttgatc aaaaaatatt 87240  
 gaaataaatc cagctggcaa aggtgagatt gcaaaaaata aatgttttgc aagcgggcat 87300  
 caccatcttt tagttaatat cgaagcattg ccagagagct ttattccttt tgacaagggt 87360  
 tatttacatt ttggaggagg tcaaaactgaa acaattcttg atcttgatcc tggaacctat 87420  
 tctctccaac taattcttgg atcttatgtg cataattcaa aaatgcaggt aaataacttt 87480  
 aaaggtcaag gaccctttt atcagaaaaa ataacaatta cagttaatta gagattagac 87540  
 cagttataac ttatatctaga tgttcatcta agttattagc aattatattt gactagaaaa 87600  
 taggaccaga attatcgtca taatcatcac caactacatt cacctttctt attatcccaa 87660  
 caagaccatg caacatagac caaagagtta tacatttaaa ggcaatcact tcctcaggct 87720

09847513 - 0000001

cgtctgctag attagcaaaa ctttttctca tattatcgta tgttccattt gcagatttaa 87780  
gaagatctgg gtaatcggca aagttaccaaa cagctgttcc aaacatcaaaa tcatatgtat 87840  
gtgcattttt taaaccaaac cctatatatt tgcttgcatt agtaacaagt tttttctttg 87900  
taattttttt tggattctca aaaaaaaciaa gctcattaag ttttttaaaag ccaacgggtg 87960  
caacagcggc atatacacat tcttttgttt caaaatgcct ataaggagca gtttgagaaa 88020  
caccactttc ctttgcaagg gatctaatac ttagtttagt gtaaccatct ctatcacaaa 88080  
gcctgcatgc gcattctata agttcttctt ttaagtttcc gtgatgatag ttattcataa 88140  
tttaatttta atgttgacac tgctaacata tatcatgtta ttatgtatac accgcataca 88200  
ttaagttgat tacatatata agtcaaatac agatatatga acataacaaa aataaataac 88260  
accgtcatag ctctttttatt aggcaatgcg tttctatcaa atcttgaggc tttagaagtt 88320  
cttgaggtta aaatgcttga tgaatatgct gtaactagag aatttccctgg aaagctcatt 88380  
ccatcagacc agtctaagct agcatttgaa atacctggaa agataaaactc tattaatgtt 88440  
gatatcggag atgagggtcat cttaggggat gaacttgcct cattggatga tagagaagct 88500  
ttagcacaaat taaatcaatc aaagggtctaaa tttgatttag ctgaacaagt actagcaaga 88560  
tatataaatc tcagagcaga tgggcatatt tctattccaag atcttgataa ggctgaatct 88620  
gatcaaatag tagctaagtc gcagtatgat ttttatagag ttaaaattga gcaaaactaag 88680  
ttactagctc cttttaatgg agttattcaa aatagatttc tagatacagg atcagtaatc 88740

09/847,513 DeLong et al.

aatgcagggtg tccaagttttt agaaatttta ggctctagca atgttgaagc aagaattttct 88800  
attccaatga actatatgaa caagggttaag attggagatg agtatgaatt tgatatcaga 88860  
ggaatatcta caaaagctac gtagagaga ctggctccca tgtcaaccgg aggctccaat 88920  
aataggtag caatttttag atttgatacc ttttttaac caggatcaat agtaaagctc 88980  
aaactaagca tcaactgagaa agcaaaaaggt acatgggttc caattaagtc actgtcccag 89040  
tccgaacaag gtatttgggc catatatacc attaacgagc acaaatagat tgtagagat 89100  
cttgtagatg ttattttattt tgaagacgaa tatgcttttg tcagcggaac acttaataat 89160  
ggtgatttag taatttttag cggagctcaa aaaattattg aaggaaaaat aataaaaaata 89220  
aatgaatggt attaattttc taatagaaaa gcctaggata ttattttctaa ctttagcatt 89280  
tatattactt tctggaattt cttctgggct ttcagttcca attcaagaaa accctgaact 89340  
ggctgagaga tggggagggtg ttcgtatttt tcttccccgg gcattccccag aaagaattga 89400  
aacagagata gtaaatgac tagaaatcaa acttagagaa gttgaagaaa tcgatgagct 89460  
tgaatcaatt attactcaag gttttttcaac aattgtagtt gaattaaatc aaagtgtacc 89520  
tcctatactt attgaagaga cttgggtccaa ggttcaagac aagctcaatc aaatagttat 89580  
tcctcaaggt gcagaaatat ttcttgatag aagcagtggt ccgcctatca ctgttcagta 89640  
tgctgtaacc tggaacggca gtggagatgc tccactaata atgatgtcca gactagcaag 89700  
ccagctaaaa agaaaaattaa gctcaatagg ctcatctcat caaactgcaa tttttggtga 89760  
aacagatgaa gaaattttta ttgaactaga ttcatcaaaag ctatcttcgc ttggattatc 89820

09/847,513,00000.1

atttcaagat atcgcaagtg ctattcaagc cctagatgca aaaaaaccta ttggtgtatc 89880  
 ctcaaaacaac aattctgagc ttttatatag actcaagat aatatacaga gcattcaaaa 89940  
 actctcagaa atacctatca aggttattaa taaatcagag atcatacagc tagatgatgt 90000  
 ggcatttatt tcaaaaaatcc cggtttctcc tattgaagac atattcttgt ttaatggaaa 90060  
 tgtagttatt tctgttgctg gaaccggatc attttctcaa agagtccatg attatgtaga 90120  
 acgcgcaaca attgttgtag atgagatgag agaaactctg ccgactgaga tcactataga 90180  
 tttagtttat gacgaatctg cttacacaac taaaaaattt aatgagcttg taaaaagttt 90240  
 ttcattagca atattttttg ttttagcttt aagtcttttt tttcttggaa ttagatcagc 90300  
 aataaattgtt actcttatcc tgccattttc tatttgcttg gttatgattg gttgtaggtt 90360  
 tataggctta cccttgcata tgacatctat cactggaatt attattgcac taggattgct 90420  
 tatagataat gggattattg ttgttgaaga ttataagaat agaagagcat ctggattaaa 90480  
 tatcaatgat tcaatttcac aaggactaaa aaacttatgg gctccattag ctgctgcaac 90540  
 agcaacaacc gtcttctctt ttcttcctat tgttactgga gaagggtcga gtattgaatt 90600  
 cgtaggcggc atggcaatga cagtcattat gtctataaca tcttcattag ttttggcggtt 90660  
 attaatgggt ccagttctga tgagttatat ggaaaaaaatt ccgttcttta aggatgtgga 90720  
 tattagcaag gaagggata gaaatgaaaa aatccttaata aaatataggg ccttttttaaa 90780  
 ctgggcgttc ttagttccta gaagagcaat catgatatcg cttgcattgc ctgttctagg 90840

09/847,513 DeLong et al. Corrected Sequence Listing (August 4<sup>th</sup>, 2001)

attctttctt tttaattctt tacctaaaga tttctttcct gctcaagata gagatatgtt 90900  
 tagagttaat atagaactgc cttctaacgc ctcatcactt actacaatgc agagagttaa 90960  
 ggaaaattaga gaagatatct tagatagtga tttaatttca atagaaaaaag attattcgtt 91020  
 tatcggcaga atgatgccta gagtttttgat gaatgttgtt ggtggagaag aaaaaacaagg 91080  
 atccaataat attgcgcagt ctgtattttt tgctactgat tattatgaaa tgattgaaaa 91140  
 ccttccagat ttatcaagaa gactgggttaa aaataaccct gacattatag ttaatatattga 91200  
 tagttttctg tctggccccc cggttttttt agatgttagt tatgtaattt ttggagatga 91260  
 tccagattta ctaaaatcac ttggtgagga gctagagcta attattaaca attctcctga 91320  
 tgtgagtctt acgaaatctg caacttcaaa ctcaataaacc aatgttgagt ttgaacttaa 91380  
 cagctcaaat atttcactgt ctggtcaaaa tgccaattat cttgtaaatg aaatgtttac 91440  
 tgcaaacaaat ggaatatattg ttggcactat gttggattca acaaaagaaa taccagtcag 91500  
 gctgaaaggg ctgtctaata aaaacaatat tacgggaaat actagtttta taacaatgcc 91560  
 ctctcaagggt gggttttgagt attttgatag tttcggaaaa agctcactaa caaacaaatc 91620  
 gtcaacaatt actaggcttg atggccaaaag acaaaatgat gttgagggtt ggatttggac 91680  
 aggtacgctc ccgtctgcta ctgaaaaaatc tattaaaaaa gatgttaaaag attttgaatt 91740  
 aagattgcca ataggctatt cattaaaaa acttggcgag gctgaaagca ggggcccaatc 91800  
 tcaagcctca ttatactctt cagctttttat gtatttcatt cttataatag taggcttggt 91860  
 tatggcgctt aattctttca gagaggctgg tctaattttg tctgttgcat tcttatcaat 91920

09/847,513 DeLong et al. 89/235

tggactatcc tttcttggtt tatttatagg ccagcaaat tatggattta ttggaactat 91980  
 aagtgcaata gggttaattg gcttatcaat aaatgattca attattgtct tatctcacat 92040  
 aaaagaagag gctgagaaga aatcactaac caaagctgag ctgttgtgaag ttgttatcag 92100  
 gtctacacgt catataaatca ctacctcttt gacaacacct ggtggttttc ttccacttat 92160  
 ttttgcaagt gtattcttca aaccgcttgc ttgggcaatg agcattggag tattaggcgc 92220  
 gactattaca gccttattat atattcctgc aatgtttatg ataatgagaa aggttaagta 92280  
 ctagaacaac tttccgagca ataatttctt cctagctttt ttacaataag tgattcatga 92340  
 gtataggctc cgcaagaatc acatttaatc atagagttat caatctttt tttatgaggt 92400  
 cctgatatga gagatctgaa aaagaatacc gcagttatta tcaatataaa tatgaccaat 92460  
 ggaatcaaaa gtaatatcga ttttaataag aacatttttt agtcactaga tgctggctta 92520  
 ctatttctta taaaattcca agaccagcca gacttatctt cagatttttt tgattgctca 92580  
 tcttggtagt taatagaaat tattttttta gtatcttcaa gcaattcaat atatcccaaa 92640  
 gattcatatg aagcctctaa gatcttaaga gctctataat ttctactaga gttaggaata 92700  
 ttttcaatta cataatttgc tcttcttatt gctgctatat gcgcatcaac actaacgtaa 92760  
 taatcagctg ccgcaagctc atttcttgca atcatatttc ttaaatagat atttctttgt 92820  
 ttagcatatg tagaataatg gctatcagga aatcttggtta agaattcagt tagttctgaa 92880  
 aatgattcct tagctcctga gatatctcta tttgatagat cagtatctgt cattcttaca 92940

09/847,513 DeLong et al.



ataaagctat tatctctcgt atagctagaa aggcctttca taaaatatgc ataatacaata 93000  
 ttggatgcc taggatgaag tcttataaat ttttctgctg cagcatgcgc agcctcagtt 93060  
 tcagcattca taaattgagc ataaataagc tctacttgcg cctgttcagc atatttgcca 93120  
 aaaggatatac ttgattcaat tgcttctaataaat gaatcaatag caccaaaata attttttcca 93180  
 gccatccttc ttggggcttg atcgtaataa atttttttcag gctgttctat ttctggggcca 93240  
 tcagaattac aactaaaccaa taacagaggt acaattggca atactataaa tagtttttaaa 93300  
 attagtttat ctttcattat ttgcacattc tacctgcatt atcgtaattt aagcttgaaa 93360  
 acatctatct agattaattt ttttataata agttcatatg attgttataaa atgttccaaa 93420  
 tgatctatct tcaatgagc tagacaaagc tacagcagaa atgttttacag attattccag 93480  
 aactcagata aaaaaatgga tagaagaagg cagagtctct ataaatggag aggtatccca 93540  
 gccaaagat aaagtttatg agaattgatca gattgaatta agccctaaag aagaacaaaa 93600  
 agtatcatgg gaagctcaag atatatgattt tgaaattcac tttgaaaaatg aagattttat 93660  
 tataaattaat aaacctgctg gtttaataat gcatccaggt tctgggtgct atgatggaac 93720  
 tctcgcaaat gggctcattt ataaatttcc agaattgatc aatattccaa ggtcggggaat 93780  
 tgttcatcga ttagacaaaag atacttctgg cattctgctc gtagcaagaa atgagtcttt 93840  
 taggaacttt tttattaatg aaatgcagga gaggagagtt gtaaaaaaat atacgtctat 93900  
 tgttattggt tctacactag gaagcttttc tatagaagag ccaattggaa gagataaaaa 93960  
 taatagaacc aaatgggcaa ttcgagaaga tggcaaagat gctttaacat ttgtaaagct 94020

09/847,513 DeLong et al.

taaagaaaat attggaaact actctgtgtt ggataataga atagagacag gaagaactca 94080  
 tcagattagg gtccatctat catcaaaaaa actaccaata attggagata aaacctatga 94140  
 cccaagcagg tctattgcaa gagatacccc tgaagagcta attaatatta tccgaggttt 94200  
 tccaaggcag gcattacatg caacacacct ctcatccaat gaccaaaaaa caaataatat 94260  
 tttttctttt gatattccca ctccaaatga tatggaggaa ctacttctag aattaagaaa 94320  
 attgatctaa tagtaactaa aaacttgttt tttgattaat aaaataatat aaaccttatt 94380  
 cctaagagtt ttttggttaag aaattgaaat tatctggcgc agacatgcta atgcaagcac 94440  
 ttcatgatga aggtgttgag ctaatctttg gctacccagg tggagccgcg cttcatatct 94500  
 atgatgcaat ttttagacaa gataaaatag atcatatttt agtaaggcat gagcaagggtg 94560  
 caaccatgc agcagatgga tattcaaggg cgacaggtaa gccaggagtt gtcttagtca 94620  
 cttctggacc tggtgcaaca aatgctatta caggaatcgc gactgcattt atggattcca 94680  
 taccaatggt agttatttca gggcagggtg ctagccattt aataggtagt gatgcttttc 94740  
 aagaaaactga tatgattggt gtttcaagac caattgttaa gcatagctat acagttttta 94800  
 atgctgaaga aatacctaag ataattaaag aagcttttta tgcgcaact tcaggcagac 94860  
 ctggacctgt tgttatagat atcccaaaag acatgacagc tccggataat cttttttgatt 94920  
 actcgtatcc tgaagaagcc aagataagat catacaatcc tccgattgag ccagaaaaaa 94980  
 atcaaataga tagagcagtc gaagctatat tgatatcaaa aaaaccagtt atatatgctg 95040

09847513\_0000001

gtggtggggc aattgctagt aatgccgaaa aagaattact tgaacttaat gaaattattg 95100  
 atgctcctgt taaaaatact ttaatgggat tgggtatttta ccctgctagt catcatagat 95160  
 ttcttgggat gttagggatg catggaacat atcaggcaaa tatggcaatg cataatgcag 95220  
 acttaataat tgctattggc gccagatttg atgacaggat taccaataaa ccatcaaagt 95280  
 ttgcacctaa tgccaaagtg gttcatctag atgttgatca ctcatctgta tcaaaaaatta 95340  
 tagaagcaaa tgtagctgtt tttaggcaag taaaaaattc cttaaaaaatta ataaaaagaaa 95400  
 ctcttgaaaa aaaattagac tcttacgatt ctttcgctct tcagccttgg cacgatcaga 95460  
 taaaagaatg gaaatcacta catggtttta attatgagct ttataaagat gaatctgatg 95520  
 atcatcccat ttaccgccag gctgtagtcc agcatgtcca tgagattaca aatgggggaag 95580  
 catatgtgac ttccgatgtt ggtcagcatc aaatgtttgc tgctcaatat tatcattttg 95640  
 ataagcctag aagatggatc aattctggtg gtctaggaac tatgggtttt ggtttgccag 95700  
 cagcaatggg tgtaaaactc gcttttccaa aagatgaggt tgtttgcatt actggtgagg 95760  
 gtagtatcca aatgtgcac ccaagaattgt ctacatgtct tcaatataat ctcccaataa 95820  
 aaataattaa tatcaataac gaagctcttg gstatggttaa acaatggcag gatatgaatt 95880  
 atggaggaag gcactctgaa agtacctatc aaaactcgtt accggatttt ataaaaactgg 95940  
 ctgaatcata tggatcata ggaattaaaa ttacaaaaaa ttctgattta agtgaaggct 96000  
 taaaaaaagc ttttgaaatg aaagataaac ttgtctttgt tgatatattat gtagatcctt 96060  
 cagagcatgt ttatccaatg caggttgcaa atggcagcct agaaaaatatg tggctatcaa 96120

09/847,513, 0000001

aggatgaaca aacatgatta aaagaaaact aattttaatt atgaaaaata aaccaggagc 96180  
 tctagtaaga gtagtggac tgtttcatca aagaggctac aacattgaaa cctttcatgt 96240  
 agatactgtt aaagactttt ctacttacia atcgatatgt aaaaaaaact tgaaccacaa 96300  
 tttgaggata atcaaatatc tagactgacc atagaaaaca tggtttcaga tgaccttatg 96360  
 aggcagattt tgagacagct caataaatta atagatgtta tagctgtaag caatgaagag 96420  
 acaacctatt taaaaggagt attattagat gaaaatttat tatgacgaag atgcaaacat 96480  
 agaaaattatt aaagggatga atgtctcaat aattggctat ggaatctcaag ggaatgctca 96540  
 tgcaataaat cttcatgaat ctggtgtaag tgttactgta ggtttaagag aagggctcttc 96600  
 ttcatgggca aaggcagaag aagcaggctt aaaaattcaa acagttgctg attcggtaat 96660  
 ccaagcagat ttggttatga ttttggcacc agatgaattc caaaaaata tatatgaaac 96720  
 cgaaatcaag ccaaaacttaa aaaccagtgc aattccttgca tttgcacatg gctttaatat 96780  
 tcattttgaa aaaatagtgc ctgaagcaac taacagcgtta attatgattg caccacaaagg 96840  
 tccaggccat actgttagaa gtactttatac caatggtgga ggcgttccat ctctcatagc 96900  
 tatatatgaa gatgcttttaa gtgatgaaga ttattcagca aaagatgtag ctctatctta 96960  
 tgcaaaaagca aatggcggca cgagggctgg tgttcttgaa acatctttta aagaagaaac 97020  
 agaaacagat ttatttggcg acaaacgagt tttatgtggc gggcttaccg ctttaattaa 97080  
 agctggggtt gaaactctag ttgaggcagg ctacagtga gagatggcat attttgaatg 97140

ccttcacatgaa aaaaaaactaa tcacagactt aattcaagaa ggtggcattg ctaatatgca 97200  
ttactcaata tcaaataactg ctgagtatgg tgattatgtg agtggaccca aagtaattac 97260  
tagcgatacc aaaaaagcaa tgaaaggaat actagaaaaat atacagtcag gaaaaattcgc 97320  
agatgacttc ttgaatgact gtcgacaaaag caatgatggc actggcggac ctgttatgaa 97380  
aagcaataga gaagctacaa aaattcatcc aatagagtcg gtagggggctg agctaaggtc 97440  
taaaatgaag ttcttaaatt cacaaaaaatt ggtggataaa gaaattaatt aaaattaatt 97500  
aaaaaaaaagg tatcttcttc ggaataaactt cgttaagata cgcgtccgca caaagaggcg 97560  
gttgttcttt aaaaaatattt ggttactcgt gtgggtgttc aaaatacagag aaaaaataat 97620  
ttagattttt tataaaaaatc acaaaaacat gatattaatt gaagagtttg atcatggctc 97680  
agattgaacg ctggcggtag gcttaacaca tgcaagtcgt gcgagaaaagt atcttcggat 97740  
atgagtagag cggcgggacgg gtgagtaacg cgtaggaatc tacctagtag aaggggatatag 97800  
cccggggaaa cccggattaa taccgtatac ctcttcggg agaaagaagg cctctctttg 97860  
aagctttcgc tactagatga gcctgcgtaa gattagcttg ttggtgaggt aaagggtcac 97920  
caaggcgacg atcttttagct ggtctgagag gacgatacgc cacattggga ctgagacacg 97980  
gcccgactc ctacgggagg cagcagtgagg gaatatggga caatggggcg aagcctgatac 98040  
cagccatacc gcgtgtgtga agaaggcctt cgggttgtaa agcactttaa gcaggggagaa 98100  
aaagttataa gttaatacct tataaccctg atgttacctg cagaataaagc accgggctaata 98160  
tccgtgccag cagccgcggt aatacggaag gtgcaagcgt taatcggaat tactgggcgt 98220

09/847,513 : 0000001

aaagcgcgcg taggtgggtt gttaagttgg atgtgaaagc cctggggctca acctaggaac 98280  
tgcattccaaa actaactcac tagagtacga tagagggagg tagaattcat agtgtagcgg 98340  
tggaatgcgt agatattatg aagaatacca gtggcgaagg cggcctcctg gatctgtact 98400  
gacactgagg tgcgaaaagcg tgggtagcga acaggattag ataccttgggt agtccacgcc 98460  
gtaaacgatg acaactagct gttgggagac tatgtctttc agtggcgcag ctaacgcctt 98520  
aagttgtccg cctgggggagt acggccgcaa ggctaaaact caaatgaatt gacgggggacc 98580  
cgcacaaagcg gtggagcatg tggtttaatt cgatgcaacg cgaaaaacct tacctactct 98640  
tgacatactt ggaggctctt gtaatgagag tgtgcctttt ggaaccaaga tacagggtgct 98700  
gcatggctgt cgtcagctcg tgtcgtgaga tgttccgta agtcggataa cgagcgcaac 98760  
ccttaccctt atttgcacg gattcggctg ggaactataa ggggactgcc ggtgataaac 98820  
cggaggaagg tgaggacgac gtcaagtcat catggccctt acgagtaggg ctacacacgt 98880  
gctacaaatgg gagatacaga cggacgctaa gccgcgaggt ggtgctaata ctaaaaagtc 98940  
tttcgtagtc cggattggag tctgcaactc gactccatga agtcggaata gctagtaata 99000  
gcggatcagc atgccgcggt gaatacgttc tcgggtcttg tacacaccgc ccgtcacacc 99060  
atggaagtgg attgcaccag aagtagatag tctaacctta gggagggcgt ttaccacggt 99120  
gtgcttcctg actgggggtga agtcgtaaca aggtagccgt aggggaacct gtggctggat 99180  
cacctcctta acgataaata gcgtttttaa cggccacacg agtaatcaaa tattaaaaaa 99240

aagaacattt agatatgtaa aatcatgtgt atgtaatttt ctagtgtata catttatgta 99300  
tacataagat cactgcaatt aaaaagtaac atatgcattt atgtgtatgt taataaaagta 99360  
attaatatat tttattaaagt tactctcaaa aatgaagata aaacttcaaa aaaaatatgt 99420  
aacctttttt aaggttatat gatcaagtaa aggaagagca caaggcggat gccttggcag 99480  
cataaggcga tgaaggacgt aataaacctgc gataagcctc ggggagctgg taaataagct 99540  
tcgatccgag gatttccgaa tgggaaaacc caatacacat aagtgtatta tcttatactg 99600  
aatacatagg tataagaggc aaacctaggg aactgaaaca tctaagtacc tagaggaaaa 99660  
gaaatcaaca gagattccgg tagtagcggc gagcgaaacc ggaccagccc ttaagcttat 99720  
ttagtccag caaaatattc tggaaagttt agccatagta ggtgatagcc ctgtatgcga 99780  
aagactaatt taagtgaat cgagtaggtc gggacacgag aaatccttgac tgaacatggg 99840  
gggaccatcc tccaaggcta aatactctat gctgaccgat agtgaaccag taccgtgagg 99900  
gaaaggcgaa aagaaccccc gcgagggggag tgaaatagaa cctgaaaacct tgtgcttaca 99960  
agcagtcgga gcagacttgt tctgtgacgg cgtacctttt gtataatggg tcaacgactt 100020  
aatttcagta gcaagcttaa ccatttaggg taggcgtagg gaaaccgagt cttaatatagg 100080  
cgctcagttg ctggaattag acccgaaacc gggtgatcta tccatggcca gtgtgaaggt 100140  
cgagtaacat cgactggagg cgcgaaccca cttatgttga aaatgaggg gatgagctgt 100200  
ggataggagt gaaaggctaa tcaaaccccg agatagctgg ttctcttcga aaactattta 100260  
ggtagtgct cgtgtattac ttaggggggt agagcactgt ttccggctagg gggtcacccc 100320

gacttaccaa accgatgcaa actccgaata cctacaagta tgagcacggg agacagactg 100380  
cgggtgctaa cgtcccgtagt cgagagggaa acaaccacaga ctgtcagcta aggtcccaaa 100440  
ttatgattaa gtgggaaaca atgtgggaag gcacaaacag ctaggaggtt ggcttagaag 100500  
cagccatcct ttaaaagaaag cgtaaatagct cactagtcca gtcggcctgc gtggaagata 100560  
taacgggggt aaatcataaa ccgaagctac agatcttaaa tttatttaag atggtagaag 100620  
agcgttctgt aagcgggtga aggtaagctg agaggcgaac tggacgtatc agaagtgcga 100680  
atgttgacat gagtaacgat caaagaggtg aaaaacctct tcgccgaaaa accaagggtt 100740  
cctgtccaac gctaatcgag gcagggtgag gcggccccta aggcgagggc gaaagccgta 100800  
gtcgatggga aacagggttaa tattcctgta cttttataa ctgcgatggg gtgacggaga 100860  
aggttagact agcacggcga cggttgtccg tgttcaaggt tgtaggctgg tgttctaggt 100920  
aaatccggaa cgctaaggct gagaactgat aacgaccact ctacgagtgg gaagtagtcg 100980  
ataccatgct tccaggaaaa acctctaagc ttcagggttat aagaaaacct accctaaacc 101040  
gacacaggtg gttaggtcga gtagaccaag gtgtttgaga gaactatggt gaaggaaacta 101100  
ggcaaaaatag caccgtaact tcgggagaag gtgcgccgag tttgggtgatg agacttgctc 101160  
tctaagctga acgtggtcga agataccagg tggctgcgac tgtttactaa aaacatagca 101220  
ctctgcaaac tcgtaagagg aagtataggg tgtgacgcct gcccggtgcc ggaagggttaa 101280  
ttgatggggt tagcttatgc gaagctcttg atcgaagccc cggtaaacgg cggccgtaac 101340

09847513.000001



tataacggtc ctaaggtagc gaaattcctt gtcgggtaag ttccgacctg cacgaatggc 101400  
gtaacgatgg ccacactgtc tccaccatag actcagtga attgaaatcg ctgttaagat 101460  
gcagtgtacc cgcagctaga cggaaagacc ccgtgcacct ttactatagg ttgcgactgg 101520  
actttgacct tactttgtga ggataggtgg gagactttga agcagagacg ccagtccttg 101580  
tggagtcac cttgaaatac cacccttgta agattgaagt tctaacctag gtccattatc 101640  
tggatcaggg acagtgcgtg ctgggtagtt tgactggggc ggtctcctcc taaagagtaa 101700  
cggaggagta cgaaggatc cttatcacgg tcggacatcg tgaggtaagt ataaaggcag 101760  
aaggatgctt gactgcgaga tcgacggatc gagcaggtag gaaactaggt cttagtgatc 101820  
cggtggttct gaatggaagg gccatcgctc aacggataaa aggtacgccg gggataacag 101880  
gctgataccg cccaagagtt catatcgacg gcggtgtttg gcacctcgat gtcgggctcat 101940  
cacatcctgg ggctggagca ggtcccaagg gtatggctgt tcgccattta aagtgggtacg 102000  
cgagctgggt ttagaacgtc gtgagacagt tcggtcccta tctgctgtgg gcgttttgag 102060  
atttgagga agctgattct agtacgagag gaccgaattg gacgaacctc tgggtgttccg 102120  
gttgtcacgc cagtggcatt gccgggtagc tatgttcgga aaggataacc gctgaaaagca 102180  
tataagcggg aagcctctcc caagattaaa tctcccagag actttatgtc tcctaaagag 102240  
tcgtcataga ctatgacgtt gataggcaag atgtgtaagc gctgcgaggc gttgagctaa 102300  
cttgtactaa taactcgtga ggcttgatca tgtaacctta agcaaggctc ataatttgag 102360  
taaaacattg tagtgagaat taaaaataa aaagttacat accagtttgc ctgatgacaa 102420

09/847,513, 0000001

tagcaacttg gaaccacctg atcccatctc gaactcagaa gtgaaacgag ttaacgcaa 102480  
tggtagtga gggctctccct gtgtgagagt aggaaatcgt caggcttttt tctttaaggc 102540  
ttccagttta ctggaagcct ttttttttat ctcaagtata ataccagat gattattgga 102600  
ttaacaggag gcattgggtc tggcaaatct gccgctgcag acttctttat tgatttaggt 102660  
atatacgtct tagatgcaga tcaagttgct aaagaagcct tatctacaaa ttctcctgga 102720  
tatactgatt ttattttctca atttgggtgaa gtgtatttaa ataataatcg tgaggttgat 102780  
aggctgaaat tgcgcgaact tattttttca aatcctttca aaaaaaggga tcttgagaat 102840  
attattcatc ccatagttcg gtctgtctatt agtaatttta ttattacatc aacatcacca 102900  
tattctattg ttatggtgcc actcattttt gaaacaaaatt catataaaaa ttacgataag 102960  
attattactg ttgactgtga tttagaactt cagatagtaa gggcctcaag tagagatgct 103020  
caaaataaat cgcagattaa gaatatattatt aataagcaag cctctagaga ggagaggcta 103080  
agtattttctg atgacgtact tatcaataac agcacccctat ctgatctaaa aaaacaagtt 103140  
aatgtttttac atactaaata tatggagtta ttaaatgagt agttgcccta gatgtgaaaa 103200  
acctgtcaaa cttttctactg acaatatatta tagacctttc tgctctgaga aatgcaaaact 103260  
tatagatttt ggtgattggg ctaatgaaga taataaaatt tcaagaccaa ttcaatctga 103320  
agattttttac gaagattaaa tttaagaaag tctccattca ccactttcaa tcatcggttt 103380  
ggcttttttta tacttcatctt ctgcgtatc ttgaccgtta gtaatttttta caagatcatt 103440

tctgcctaac ttaggttcgt ttcttgtcac agtacttggt ttactatag gagggttg 103500  
tttttcatt tgaagatat cagaattaat ttctctttt tccaatttta attcttgaga 103560  
actatatttg ttgtttatgc tctctagctc agattctgtc gatatttgta atgagaaaaag 103620  
tattcttata gtttcaacat ctatttccga aagcatagat tcaaacattg aatatgcctc 103680  
tcttttgaat tcatttttttg gattttttttg agcataagcc cttagaccaa cactatttct 103740  
taaatgatct atttctgata aatgctcttt ccaatggaca tcaagtactt gcagcataac 103800  
ctgcttctca agaagcaatc tattttcacc aagatcacta aacttttttg aatatttatt 103860  
ttttgcttgc agaacaattt ctctggcaat tgtatttggt acaagttttt tattactatt 103920  
aattttattg gctatatctg ttcttagtcc atagctctct tttaataaat catcgagctg 103980  
cctacttttc cattgagact ctattgattc ttcagggtaca tataaattag atattccttt 104040  
gaattgctgc tcgatgagtg actcaatggt actgctaata tcttcctctt ctagtaattg 104100  
attccttaaa gaatatatag cttgtctttg atcatttgat acatcgtcgt actctaaaaag 104160  
atttttctct gcatacaagt ttctgctctc aattcttttt tgcgcatttt ctattcctct 104220  
tgaaagcatt ttagcttcaa tatgatcatc tcccatgccc agcctttcaa aagtagccct 104280  
tcttcccatcc gaaataaaaa gtctcaagag atcgtcttct aaagataaga agaatttttga 104340  
ataaacctgga tcaccttgtc tgcctgatct ccccttaac tgattatcta ttcttcttga 104400  
ttcatgcctt tctgttccaa gtatatgaag cccaccagat tctataactt ttttattatt 104460  
ttctttccac tctatatcgc ttgtgtcttc ttttttctcg ccaagaacta tatccgtccc 104520

09847513.0000001

ccttcctgcc atattcgtag caatagttac cattccagggt ttacctgcat tggcaattat 104580  
ctctgcttct ttttcgtgat gcttggcatt taaaatttga tgaggatttt ttttgttatt 104640  
taaatatgct gatacttctt ctgaagattc aaccgaaact gttccaacaa gaatgggaga 104700  
agattttttt cttaattggt caattttctt aattagagct ttatatctcg attctgttgt 104760  
aagaaatact aagtcattaa gatcagctct agccatagga acatttggtg ggatgatgat 104820  
gacatttagg ccatagatatt gactaaactc tactgcttca gtatctgctg ttcctgtcat 104880  
cccagaaagt tttttaata atctaaaaaa gtttttggat gtggtggatg ctagtgtttg 104940  
agactctctt tggatagcaa ctttttcttt gcattccagt gcctggtgaa caccttcaact 105000  
cattcttctt ccgggcattg ttctacctgt atgctcatca atcaaaaaga cctcaccgtt 105060  
cctaaccaaa taatccacat tcttttttaa taagaagctt gctctaagtg ttgcttgaac 105120  
aaatttcata atttttaaat tagaaaacaga gtaagcccat ctgaggctcc aagccgattc 105180  
cagc 105184

<210> 2

<211> 29

<212> DNA

<213> pcr primer

<220>

<221> misc\_feature

<222> (1)..(29)

<223> Oligonucleotide primer for amplifying the proteorhodopsin gene. F

orward primer

<400> 2 29  
accatgggta aattattact gatattagg

<210> 3  
<211> 24  
<212> DNA  
<213> pcr primer  
  
<220>  
<221> misc\_feature  
<222> (1)..(24)  
<223> Oligonucleotide primer for amplifying the proteorhodopsin gene.  
Reverse primer

<400> 3 24  
agcattagaa gattctttaa cagc

<210> 4  
<211> 750  
<212> DNA  
<213> Naturally occurring gamma proteobacterium  
  
<220>  
<221> CDS  
<222> (1)..(750)  
<223> Light-driven proton pump that is active when expressed in E. coli  
, with retinal. An additional three nucleotides are incorporated  
to native sequence (31A08) via pcr primer (DNA residues 4-6, ggt  
) , adding a new restriction site for cloning/expression

<400> 4  
 atg ggt aaa tta tta ctg ata tta ggt agt gtt att gca ctt cct aca 48  
 Met Gly Lys Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr 15  
 1 5 10  
  
 ttt gct gca ggt ggt ggt gac ctt gat gct agt gat tac act ggt gtt 96  
 Phe Ala Ala Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val 30  
 20 25  
  
 tct ttt tgg tta gtt act gct gct tta tta gca tct act gta ttt ttc 144  
 Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe 45  
 35 40  
  
 ttt gtt gaa aga gat aga gtt tct gca aaa tgg aaa aca tca tta act 192  
 Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr 60  
 50 55  
  
 gta tct ggt ctt gtt act ggt att gct ttc tgg cat tac atg tac atg 240  
 Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met 80  
 65 70 75  
  
 aga ggg gta tgg att gaa act ggt gat tcg cca act gta ttt aga tac 288  
 Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr 95  
 85 90  
  
 att gat tgg tta cta aca gtt cct cta tta ata tgt gaa ttc tac tta 336  
 Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu 110  
 100 105  
  
 att ctt gct gct gca act aat gtt gct gga tca tta ttt aag aaa tta 384  
 Ile Leu Ala Ala Thr Asn Val Ala Gly Ser Leu Phe Lys Lys Leu 125  
 115 120

09/847,513 DeLong et al.

cta gtt ggt tct ctt gtt atg ctt gtg ttt ggt tac atg ggt gaa gca  
 Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala  
 130 135 140 432  
  
 gga atc atg gct gca tgg cct gca ttc att att ggg tgt tta gct tgg  
 Gly Ile Met Ala Ala Trp Pro Ala Phe Ile Ile Gly Cys Leu Ala Trp  
 145 150 155 480  
  
 gta tac atg att tat gaa tta tgg gct gga gaa gga aaa tct gca tgt  
 Val Tyr Met Ile Tyr Glu Leu Trp Ala Gly Glu Gly Lys Ser Ala Cys  
 165 170 175 528  
  
 aat act gca agt cct gct gct gtg caa tca gct tac aac aca atg atg tat  
 Asn Thr Ala Ser Pro Ala Val Gln Ser Ala Tyr Asn Thr Met Met Tyr  
 180 185 190 576  
  
 att atc atc ttt ggt tgg gcg att tat cct gta ggt tat ttc aca ggt  
 Ile Ile Ile Phe Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly  
 195 200 205 624  
  
 tac ctg atg ggt gac gga tca gct ctt aac tta aac ctt atc tat  
 Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Ile Tyr  
 210 215 220 672  
  
 aac ctt gct gac ttt gtt aac aag att cta ttt ggt tta att ata tgg  
 Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile Trp  
 225 230 235 720  
  
 aat gtt gct gtt aaa gaa tct tct aat gct  
 Asn Val Ala Val Lys Glu Ser Ser Asn Ala  
 245 250 750

<210> 5  
 <211> 250

09/847,513 .0000001

<212> PRT  
<213> Naturally occurring gamma proteobacterium

<400> 5

Met Gly Lys Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr  
1 5 10 15

Phe Ala Ala Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val  
20 25 30

Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe  
35 40 45

Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
50 55 60

Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met  
65 70 75 80

Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr  
85 90 95

Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu  
100 105 110

Ile Leu Ala Ala Ala Thr Asn Val Ala Gly Ser Leu Phe Lys Lys Leu  
115 120 125



Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala  
130 135 140

Gly Ile Met Ala Ala Trp Pro Ala Phe Ile Ile Gly Cys Leu Ala Trp  
145 150 155 160

Val Tyr Met Ile Tyr Glu Leu Trp Ala Gly Glu Lys Ser Ala Cys  
165 170 175

Asn Thr Ala Ser Pro Ala Val Gln Ser Ala Tyr Asn Thr Met Met Tyr  
180 185 190

Ile Ile Ile Phe Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly  
195 200 205

Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Ile Tyr  
210 215 220

Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile Trp  
225 230 235 240

Asn Val Ala Val Lys Glu Ser Ser Asn Ala  
245 250

<210> 6

09/847,513, 0000001

<211> 747  
 <212> DNA  
 <213> Naturally occurring gamma proteobacterium  
  
 <220>  
 <221> CDS  
 <222> (1)..(747)  
 <223> Native proteorhodopsin DNA sequence from BAC clone 31A08  
  
 <300>  
 <301> Beja, O., Aravind, L., Koonin, E.V., Suzuki, M.T., Hadd, A., Nguyen, L.P.,  
 Jovanovich, S.B., Gates, C.M., Feldman, R.A., Spudich, J.L., Spudich, E.N. and DeLong, E.F.  
 <302> Bacterial rhodopsin: evidence for a new type of phototrophy in the sea  
 <303> Science  
 <304> 289  
 <305> 5486  
 <306> 1902-1906  
 <307> 2000-09-15  
 <308> AAG10475  
 <309> 2000-06-15  
 <313> (1)..(747)  
  
 <400> 6 48  
 atg aaa tta tta ctg ata tta ggt agt gtt att gca ctt cct aca ttt  
 Met Lys Leu Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr Phe 15  
 1  
  
 gct gca ggt ggt ggt gac ctt gat gct agt gat tac act ggt gtt tct 96  
 Ala Ala Gly Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val Ser 30  
 20 25  
  
 ttt tgg tta gtt act gct gct tta tta gca tct act gta ttt ttc ttt 144  
 Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe Phe 45  
 35 40

gtt gaa aga gat aga gtt tct gca aaa tgg aaa aca tca tta act gta 192  
 Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr Val  
 50 55 60  
 tct ggt ctt gtt act ggt att gct ttc tgg cat tac atg tac atg aga 240  
 Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met Arg  
 65 70 75 80  
 ggg gta tgg att gaa act ggt gat tgc cca act gta ttt aga tac att 288  
 Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr Ile  
 85 90  
 gat tgg tta cta aca gtt cct cta tta ata tgt gaa ttc tac tta att 336  
 Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu Ile  
 100 105 110  
 ctt gct gct gca act aat gtt gct gga tca tta ttt aag aaa tta cta 384  
 Leu Ala Ala Ala Thr Asn Val Ala Gly Ser Leu Phe Lys Lys Leu Leu  
 115 120 125  
 gtt ggt tct ctt gtt atg ctt gtg ttt ggt tac atg ggt gaa gca gga 432  
 Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala Gly  
 130 135 140  
 atc atg gct gca tgg cct gca ttc att att ggg tgt tta gct tgg gta 480  
 Ile Met Ala Ala Trp Pro Ala Phe Ile Ile Gly Cys Leu Ala Trp Val  
 145 150 155 160  
 tac atg att tat gaa tta tgg gct gga gaa gga aaa tct gca tgt aat 528  
 Tyr Met Ile Tyr Glu Leu Trp Ala Gly Glu Gly Lys Ser Ala Cys Asn  
 165 170 175  
 act gca agt cct gct gtg caa tca gct tac aac aca atg atg tat att 576  
 Thr Ala Ser Pro Ala Val Gln Ser Ala Tyr Asn Thr Met Met Tyr Ile

180 185 190  
 atc atc ttt ggt tgg gcg att tat cct gta ggt tat ttc aca ggt tac 624  
 ile ile phe gly trp ala ile tyr pro val gly tyr phe thr gly tyr  
 195 200  
 ctg atg ggt gac ggt gga tca gct ctt aac tta aac ctt atc tat aac 672  
 leu met gly asp gly gly ser ala leu asn leu asn leu ile tyr asn  
 210 215 220  
 ctt gct gac ttt gtt aac aag att cta ttt ggt tta att ata tgg aat 720  
 leu ala asp phe val asn lys ile leu phe gly leu ile ile trp asn  
 225 230 235 240  
 gtt gct gtt aaa gaa tct tct aat gct 747  
 val ala val lys glu ser ser asn ala 245

<210> 7  
 <211> 249  
 <212> PRT  
 <213> Naturally occurring gamma proteobacterium

<400> 7  
 Met Lys Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr Phe  
 1 5 10 15

Ala Ala Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val Ser  
 20 25 30

phe trp leu val thr ala ala leu leu ala ser thr val phe phe phe

35

40

45

Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr Val  
50 55 60

Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met Arg  
65 70 75 80

Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr Ile  
85 90 95

Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu Ile  
100 105 110

Leu Ala Ala Thr Asn Val Ala Gly Ser Leu Phe Lys Lys Leu Leu  
115 120 125

Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala Gly  
130 135 140

Ile Met Ala Ala Trp Pro Ala Phe Ile Ile Gly Cys Leu Ala Trp Val  
145 150 155 160

Tyr Met Ile Tyr Glu Leu Trp Ala Gly Glu Gly Lys Ser Ala Cys Asn  
165 170 175

09847513.000001

Thr Ala Ser Pro Ala Val Gln Ser Ala Tyr Asn Thr Met Tyr Ile  
 180 185 190

Ile Ile Phe Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly Tyr  
 195 200 205

Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Ile Tyr Asn  
 210 215 220

Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile Trp Asn  
 225 230 235 240

Val Ala Val Lys Glu Ser Ser Asn Ala  
 245

<210> 8  
 <211> 750  
 <212> DNA  
 <213> Naturally occurring gamma proteobacterium  
 <220>  
 <221> CDS  
 <222> (1)..(750)  
 <223> proteorhodopsin variant from clone EBAC40

<400> 8  
 atg ggt aaa tta tta ctg ata tta ggt agt gtt att gca ctt cct aca  
 Met Gly Lys Leu Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr  
 1 5 10 15

48

09847513.0000001



145	150	155	160	
gta tac atg att tat gaa cta tgg gct gga gaa ggc aag gct gca tgt				528
Val Tyr Met Ile Tyr Glu Leu Trp Ala Gly Glu Lys Ala Ala Cys				
165	170		175	
aat act gca agt cct gct gct gtg caa tca gct tac aac aca atg atg tat				576
Asn Thr Ala Ser Pro Ala Val Gln Ser Ala Tyr Asn Thr Met Met Tyr				
180	185	190		
ata atc atc ttt ggt tgg gca att tat cct gta ggt tat ttc aca ggt				624
Ile Ile Ile Phe Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly				
195	200	205		
tac cta atg ggt gac ggt gga tca gct ctt aac tta aac ctt atc tat				672
Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Ile Tyr				
210	215	220		
gac ctt gct gac ttt gtt aac aag att cta ttt ggt tta att ata tgg				720
Asp Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Trp				
225	230	235	240	
aat gtt gct gtt aaa gaa tct tct aat gct				750
Asn Val Ala Val Lys Glu Ser Ser Asn Ala				
245	250			
<210> 9				
<211> 250				
<212> PRT				
<213> Naturally occurring gamma proteobacterium				
<400> 9				
Met Gly Lys Leu Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr				



1	5	10	15
Phe	Ala	Gly	Gly
Ala	Gly	Gly	Asp
Ala	Leu	Asp	Ala
20	25	Ser	Thr
30	35	Val	Thr
40	45	Leu	Gly
50	55	Ala	Val
60	65	Ser	Phe
70	75	Thr	Val
80	85	Ala	Thr
90	95	Leu	Val
100	105	Val	Leu
110	115	Thr	Val
120	125	Arg	Thr
130	135	Glu	Val
140	145	Thr	Leu
150	155	Val	Val
160	165	Leu	Val
170	175	Val	Val
180	185	Val	Val
190	195	Val	Val
200	205	Val	Val
210	215	Val	Val
220	225	Val	Val
230	235	Val	Val
240	245	Val	Val
250	255	Val	Val
260	265	Val	Val
270	275	Val	Val
280	285	Val	Val
290	295	Val	Val
300	305	Val	Val
310	315	Val	Val
320	325	Val	Val
330	335	Val	Val
340	345	Val	Val
350	355	Val	Val
360	365	Val	Val
370	375	Val	Val
380	385	Val	Val
390	395	Val	Val
400	405	Val	Val
410	415	Val	Val
420	425	Val	Val
430	435	Val	Val
440	445	Val	Val
450	455	Val	Val
460	465	Val	Val
470	475	Val	Val
480	485	Val	Val
490	495	Val	Val
500	505	Val	Val
510	515	Val	Val
520	525	Val	Val
530	535	Val	Val
540	545	Val	Val
550	555	Val	Val
560	565	Val	Val
570	575	Val	Val
580	585	Val	Val
590	595	Val	Val
600	605	Val	Val
610	615	Val	Val
620	625	Val	Val
630	635	Val	Val
640	645	Val	Val
650	655	Val	Val
660	665	Val	Val
670	675	Val	Val
680	685	Val	Val
690	695	Val	Val
700	705	Val	Val
710	715	Val	Val
720	725	Val	Val
730	735	Val	Val
740	745	Val	Val
750	755	Val	Val
760	765	Val	Val
770	775	Val	Val
780	785	Val	Val
790	795	Val	Val
800	805	Val	Val
810	815	Val	Val
820	825	Val	Val
830	835	Val	Val
840	845	Val	Val
850	855	Val	Val
860	865	Val	Val
870	875	Val	Val
880	885	Val	Val
890	895	Val	Val
900	905	Val	Val
910	915	Val	Val
920	925	Val	Val
930	935	Val	Val
940	945	Val	Val
950	955	Val	Val
960	965	Val	Val
970	975	Val	Val
980	985	Val	Val
990	995	Val	Val
1000	1005	Val	Val

Gly Ile Met Asn Ala Trp Gly Ala Phe Val Ile Gly Cys Leu Ala Trp  
 145 150 155 160

Val Tyr Met Ile Tyr Glu Leu Trp Ala Gly Glu Gly Lys Ala Ala Cys  
 165 170 175

Asn Thr Ala Ser Pro Ala Val Gln Ser Ala Tyr Asn Thr Met Met Tyr  
 180 185 190

Ile Ile Ile Phe Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly  
 195 200 205

Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Ile Tyr  
 210 215 220

Asp Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile Trp  
 225 230 235 240

Asn Val Ala Val Lys Glu Ser Ser Asn Ala  
 245 250

<210> 10  
 <211> 750  
 <212> DNA

<213> Naturally occurring gamma proteobacterium

<220>  
 <221> CDS

<222> (1)..(750)  
 <223> proteorhodopsin variant from clone EBAC41

```

<400> 10
atg ggt aaa tta tta ctg ata tta ggt agt gtt att gca ctt cct aca 48
Met Gly Lys Leu Leu 5
1
ttt gct gca ggt ggt ggt gac ctt gat gct agt gat tac act ggt gtt 96
Phe Ala Ala Gly Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val
20 25 30
tct ttt tgg tta gct act gct gct tta tta gca tct act gta ttt ttc 144
Ser Phe Trp Leu Ala Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe
35 40 45
ttt gtt gaa aga gat aga gtt tct gca aaa tgg aaa aca tca tta act 192
Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr
50 55 60
gta tct ggt ctt gtt act ggt att gct ttc tgg cat tac atg tac atg 240
Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met
65 70 75
aga ggg gta tgg att gaa act ggt gat tcg cca act gta ttt aga tac 288
Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr
85 90 95
att gat tgg tta cta aca gtt cct cta tta ata tgt gaa ttc tac tta 336
Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu
100 105 110
att ctt gct gct gct act aat gtt gct gga tca tta ttt aag aaa tta 384
Ile Leu Ala Ala Ala Thr Asn Val Ala Gly Ser Leu Phe Lys Lys Leu

```

09847513.0000001

115	120	125	
cta gtt ggt tct ctt gtt atg ctt gtg ttt ggt tac atg ggt gaa gca			432
Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala			
130	135	140	
gga atc atg gct gca tgg cct gca ttc att att ggg tgt tta gct tgg			480
Gly Ile Met Ala Ala Trp Pro Ala Phe Ile Ile Gly Cys Leu Ala Trp			
145	150	155	160
gta tac atg att tat gaa cta tgg gct gga gaa gga aaa tct gca tgt			528
Val Tyr Met Ile Tyr Glu Leu Trp Ala Gly Glu Gly Lys Ser Ala Cys			
165	170	175	
aat act gca agt cct gct gtg caa tca gct tac aac aca atg atg tat			576
Asn Thr Ala Ser Pro Ala Val Gln Ser Ala Tyr Asn Thr Met Met Tyr			
180	185	190	
att atc atc ttt ggt tgg gcg att tat cct gta ggt tat ttc aca ggt			624
Ile Ile Ile Phe Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly			
195	200	205	
tac ctg atg ggt gac ggt gga tca gct ctt aac tta aac ctt atc tat			672
Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Ile Tyr			
210	215	220	
aac ctt gct gat ttt gtt aac aag att cta ttt ggt tta att ata tgg			720
Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile Trp			
225	230	235	240
aat gtt gct gtt aaa gaa tct tct aat gct			750
Asn Val Ala Val Lys Glu Ser Ser Asn Ala			
			250
			245

<210>	11
<211>	250
<212>	PRT
<213>	Naturally occurring gamma proteobacterium

$\langle 400 \rangle$  11

Met Gly Lys Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr  
1 5 10 15

Phe Ala Ala Gly Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val  
20 25 30

Ser Phe Trp Leu Ala Thr Ala Leu Leu Ala Ser Thr Val Phe Phe  
35 40 45

Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
50 55 60

Val	Ser	Gly	Leu	Val	Thr	Gly	Ile	Ala	Phe	Trp	His	Tyr	Met	80
65					70					75				

Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr  
85 90 95

Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu  
100 105 110

[illegible]

120/235

<210> 12  
 <211> 750  
 <212> DNA  
 <213> Naturally occurring gamma proteobacterium  
  
 <220>  
 <221> CDS  
 <222> (1)..(750)  
 <223> Proteorhodopsin variant from clone EBAC64  
  
 <400> 12 48  
 atg ggt aaa tta tta ctg ata tta ggt agt gtt att gca ctt cct aca  
 Met Gly Lys Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr  
 1 5 10 15  
  
 ttt gct gca ggt ggc ggt gac ctt gat gct agt gat tac act ggt gtt  
 Phe Ala Ala Gly Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val  
 20 25 30 35 40 45  
  
 tct ttt tgg tta gtt aca gct gct cta tta gca tct act gta ttt ttc  
 Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe  
 50 55 60 65  
  
 ttt gtt gaa aga gat aga gtt tct gca aaa tgg aaa aca tca tta act  
 Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
 60 65 70 75 80  
  
 gta tct ggt ctt gtt act ggt att gct ttc tgg cat tac atg tac atg  
 Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met  
 80 85 90 95  
  
 aga gga gta tgg att gaa act ggt gat tcg cct act gta ttt aga tac  
 Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr  
 100 105 110 115 120 125 130 135 140 145 150 155 160 165 170 175 180 185 190 195 200 205 210 215 220 225 230 235 240 245 250 255 260 265 270 275 280 285 290 295 300 305 310 315 320 325 330 335 340 345 350 355 360 365 370 375 380 385 390 395 400 405 410 415 420 425 430 435 440 445 450 455 460 465 470 475 480 485 490 495 500 505 510 515 520 525 530 535 540 545 550 555 560 565 570 575 580 585 590 595 600 605 610 615 620 625 630 635 640 645 650 655 660 665 670 675 680 685 690 695 700 705 710 715 720 725 730 735 740 745 750 755 760 765 770 775 780 785 790 795 800 805 810 815 820 825 830 835 840 845 850 855 860 865 870 875 880 885 890 895 900 905 910 915 920 925 930 935 940 945 950 955 960 965 970 975 980 985 990 995

	85	90	95	
att gat tgg tta cta aca gtt cct tta tta ata tgt gaa ttc tac tta				336
Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu	100	105	110	
att ctt gct gca act aat gtt gcc ggc tca tta ttt aag aaa ctt				384
Ile Leu Ala Ala Thr Asn Val Ala Gly Ser Leu Phe Lys Lys Leu	115	120	125	
cta gtt ggt tct ctt gtt atg ctt gtg ttt ggt tac atg ggt gaa gca				432
Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala	130	135	140	
gga att atg gca gct tgg cct gca ttc att att ggg tgt tta gct tgg				480
Gly Ile Met Ala Ala Trp Pro Ala Phe Ile Ile Gly Cys Leu Ala Trp	145	150	155	
gta tac atg att tat gaa cta tat gct gga gaa gga aaa tct gca tgt				528
Val Tyr Met Ile Tyr Glu Leu Tyr Ala Gly Glu Gly Lys Ser Ala Cys	165	170	175	
aat act gca agt cct tcg gtt caa tca gct tac aac aca atg atg gct				576
Asn Thr Ala Ser Pro Ser Val Gln Ser Ala Tyr Asn Thr Met Met Ala	180	185	190	
atc ata gtc ttc ggt tgg gca att tat cct ata ggt tat ttc aca ggt				624
Ile Ile Val Phe Gly Trp Ala Ile Tyr Pro Ile Gly Tyr Phe Thr Gly	195	200	205	
tac cta atg ggt gac ggt gga tca gct ctt aac tta aac ctt att tat				672
Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Asn Ile Tyr	210	215	220	
aac ctt gct gac ttt gtt aac aag att cta ttt ggt tta att ata tgg				720

09/847,513 , 000000.1



Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile Trp  
225 230 235 240

aat gtt gct gtt aaa gaa tct tct aat gct 750  
Asn Val Ala Val Lys Glu Ser Ser Asn Ala  
245 250

<210> 13  
<211> 250  
<212> PRT  
<213> Naturally occurring gamma proteobacterium

<400> 13

Met Gly Lys Leu Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr  
1 5 10 15

Phe Ala Ala Gly Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val  
20 25 30

Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe  
35 40 45

Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
50 55 60

Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met  
65 70 75 80

Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr  
 85 90 95  
 Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu  
 100 105 110  
 Ile Leu Ala Ala Thr Asn Val Ala Gly Ser Leu Phe Lys Lys Leu  
 115 120 125  
 Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala  
 130 135 140  
 Gly Ile Met Ala Ala Trp Pro Ala Phe Ile Ile Gly Cys Leu Ala Trp  
 145 150 155 160  
 Val Tyr Met Ile Tyr Glu Leu Tyr Ala Gly Glu Lys Ser Ala Cys  
 165 170 175  
 Asn Thr Ala Ser Pro Ser Val Gln Ser Ala Tyr Asn Thr Met Met Ala  
 180 185 190  
 Ile Ile Val Phe Gly Trp Ala Ile Tyr Pro Ile Gly Tyr Phe Thr Gly  
 195 200 205  
 Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Ile Tyr  
 210 215 220

09/847,513 DeLong et al.

Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile Trp  
225 230 235 240

Asn Val Ala Val Lys Glu Ser Ser Asn Ala  
245 250

<210> 14  
<211> 750  
<212> DNA  
<213> Naturally occurring gamma proteobacterium

<220>  
<221> CDS  
<222> (1)..(750)  
<223> Proteorhodopsin variant from pcr clone HOT01m: GenBank# AF349978

<400> 14 48  
atg ggt aaa tta tta ctg ata tta ggt agt gtt att gca ctt cct aca  
Met Gly Lys Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr  
1 5 10 15

ttt gct gca ggt ggt ggt gac ctt gat gct agt gat tac act ggt gtt 96  
Phe Ala Ala Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val  
20 25 30

tct ttt tgg tta gtt act gct gct cta tta gca tct act gta ttt ttc 144  
Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe  
35 40 45

ttt gtt gaa aga gat aga gtt tct gca aaa tgg aaa aca tca tta act 192

09847513.0000001

Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
 50 55 60  
 gta tcg ggt ctt gtt act ggt att gct ttc tgg cat tac atg tac atg 240  
 Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met  
 65 70 75 80  
 aga ggg gta tgg att gag acc ggt gat tcg cca act gta ttt aga tac 288  
 Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr  
 85 90 95  
 att gat tgg tta cta aca gtt cct cta ttg ata tgt gaa ttc tac tta 336  
 Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu  
 100 105 110  
 att ctt gct gct gca aca aat gtt gct gct ggc ctg ttt aag aaa tta 384  
 Ile Leu Ala Ala Thr Asn Val Ala Ala Gly Leu Phe Lys Lys Leu  
 115 120 125  
 ttg gtt ggt tct ctt gtt atg ctt gtg ttt ggt tac atg ggt gag gca 432  
 Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala  
 130 135 140  
 gga att atg aac gct tgg ggt gca ttc gtt att ggg tgt tta gct tgg 480  
 Gly Ile Met Asn Ala Trp Gly Ala Phe Val Ile Gly Cys Leu Ala Trp  
 145 150 155 160  
 gta tac atg att tat gaa cta tgg gct gga gaa ggc aag gct gca tgt 528  
 Val Tyr Met Ile Tyr Glu Leu Trp Ala Gly Glu Gly Lys Ala Ala Cys  
 165 170 175  
 aat act gca agt cct gct gtg caa tca gct tac aac aca atg atg tat 576  
 Asn Thr Ala Ser Pro Ala Val Gln Ser Ala Tyr Asn Thr Met Met Tyr  
 180 185 190

09847513.0000001

ata atc atc ttt ggt tgg gca att tat cct gta ggt tat ttc aca ggt 624  
 ile ile ile phe gly trp ala ile tyr pro val gly tyr phe thr gly  
 195 200 205

tac cta atg ggt gac ggt gga tca gct ctt aac tta aac ctt atc tat 672  
 tyr leu met gly asp gly gly ser ala leu asn leu asn leu ile tyr  
 210 215 220

aac ctt gct gac ttt gtt aac aag att cta ttt ggt tta att ata tgg 720  
 asn leu ala asp phe val asn lys ile leu phe gly leu ile ile trp  
 225 230 235 240

aat gtt gct gtt aaa gaa tct tct aat gct 750  
 asn val ala val lys glu ser ser asn ala  
 245 250

<210> 15  
 <211> 250  
 <212> PRT  
 <213> Naturally occurring gamma proteobacterium

<400> 15

Met Gly Lys Leu Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr  
 1 5 10 15

Phe Ala Ala Gly Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val  
 20 25 30

Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe  
 35 40 45

09847513.0000001

Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
 50 55 60  
 Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met  
 65 70 75 80  
 Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr  
 85 90 95  
 Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu  
 100 105 110  
 Ile Leu Ala Ala Thr Asn Val Ala Ala Gly Leu Phe Lys Lys Leu  
 115 120 125  
 Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala  
 130 135 140  
 Gly Ile Met Asn Ala Trp Gly Ala Phe Val Ile Gly Cys Leu Ala Trp  
 145 150 155 160  
 Val Tyr Met Ile Tyr Glu Leu Trp Ala Gly Glu Lys Ala Ala Cys  
 165 170 175  
 Asn Thr Ala Ser Pro Ala Val Gln Ser Ala Tyr Asn Thr Met Met Tyr  
 180 185 190

09847513.0000001

Ile Ile Ile Phe Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly  
195 200 205

Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Ile Tyr  
210 215 220

Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile Trp  
225 230 235 240

Asn Val Ala Val Lys Glu Ser Ser Asn Ala  
245 250

<210> 16  
<211> 753  
<212> DNA  
<213> Naturally occurring gamma prtoeobacterium

<220>  
<221> CDS  
<222> (1)..(753)  
<223> Proteorhodopsin variant from pcr clone HOT75m1: GenBank#AF349979

<400> 16 48  
atg ggt aaa tta tta ctg ata tta ggt agt gct att gca ctt cca tca  
Met Gly Lys Leu Leu Ile Leu Gly Ser Ala Ile Ala Leu Pro Ser  
1 5 10 15

ttt gct gct gct ggt ggc gat cta gat ata agt gat act gtt ggt gtt 96  
 Phe Ala Ala Ala Gly Gly Asp Leu Asp Ile Ser Asp Thr Val Gly Val  
 20 25 30  
 tca ttc tgg ctg gtt gtt aca gct ggt atg tta gcg gca act gtg ttc ttt 144  
 Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Thr Val Phe  
 35 40 45  
 ttt gta gaa aga gac caa gtc agc gct aag tgg aaa act tca ctt gct 192  
 Phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Ala  
 50 55 60  
 gta tct ggt tta att act ggt ata gct ttt tgg cat tat ctc tat atg 240  
 Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met  
 65 70 75 80  
 aga ggt gtt tgg ata gac act ggt gat acc cca aca gta ttc aga tat 288  
 Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr  
 85 90 95  
 att gat tgg tta tta act gtt cca tta caa atg gtt gag ttc tat cta 336  
 Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Met Val Glu Phe Tyr Leu  
 100 105 110  
 att ctt gct gct tgt aca agt gtt gct gct tca tta ttt aag aag ctt 384  
 Ile Leu Ala Ala Cys Thr Ser Val Ala Ala Ser Leu Phe Lys Lys Leu  
 115 120 125  
 cta gct ggt tca tta gta atg tta ggt gct gga ttt gca ggc gaa gct 432  
 Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Phe Ala Gly Glu Ala  
 130 135 140  
 gga tta gct cct gta tta cct gct ttc att att ggt atg gct gga tgg 480  
 Gly Leu Ala Pro Val Leu Pro Ala Phe Ile Ile Gly Met Ala Gly Trp  
 145 150 155 160

09047513\_0000001



tta tac atg att tat gag cta tat atg ggt gaa ggt aag gct gct gta 528  
 Leu Tyr Met Ile Tyr Glu Leu Tyr Met Gly Glu Lys Ala Ala Val 175  
 165  
 agt act gca agt cct gct gtt aac tct gca tac aac gca atg atg atg 576  
 Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Met 190  
 180  
 att att gtt gtt gga tgg gca att tat cct gct gga tat gct gct ggt 624  
 Ile Ile Val Val Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly 205  
 195  
 tac cta atg ggt ggc gaa ggt gta tac gct tca aac tta aac ctt ata 672  
 Tyr Leu Met Gly Gly Glu Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile 220  
 210  
 tat aac ctt gcc gac ctt gtt aac aag att cta ttt ggt ttg atc att 720  
 Tyr Asn Leu Ala Asp Leu Val Asn Lys Ile Leu Phe Gly Leu Ile 240  
 225  
 tgg aat gtt gct gtt aaa gaa tct tct aat gct 753  
 Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala 250  
 245  
 <210> 17  
 <211> 251  
 <212> PRT  
 <213> Naturally occurring gamma prtoeobacterium  
 <400> 17  
 Met Gly Lys Leu Leu Leu Ile Leu Gly Ser Ala Ile Ala Leu Pro Ser 15  
 1 5 10

Phe Ala Ala Gly Gly Asp Leu Asp Ile Ser Asp Thr Val Gly Val  
 20 25 30  
 Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Ala Thr Val Phe Phe  
 35 40 45  
 Phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Ala  
 50 55 60  
 Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met  
 65 70 75 80  
 Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr  
 85 90 95  
 Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Met Val Glu Phe Tyr Leu  
 100 105 110  
 Ile Leu Ala Ala Cys Thr Ser Val Ala Ala Ser Leu Phe Lys Lys Leu  
 115 120 125  
 Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Phe Ala Gly Glu Ala  
 130 135 140  
 Gly Leu Ala Pro Val Leu Pro Ala Phe Ile Ile Gly Met Ala Gly Trp

145	150	155	160
Leu Tyr Met Ile Tyr Glu Leu Tyr Met Gly Glu Gly Lys Ala Ala Val			175
165	170		
Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Met			190
180	185		
Ile Ile Val Val Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly			205
195	200		
Tyr Leu Met Gly Gly Glu Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile			220
210	215		
Tyr Asn Leu Ala Asp Leu Val Asn Lys Ile Leu Phe Gly Leu Ile Ile			235
225	230		
Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala			250
	245		
<210> 18			
<211> 753			
<212> DNA			
<213> Naturally occurring gamma proteobacterium			
<220>			
<221> CDS			
<222> (1)..(753)			

**SECRET**

<223> Proteorhodopsin variant from pcr clone HOT75m3; GenBank#AF349980

```
<400> 18
atg ggt aaa tta tta ctg ata tta ggt agt gct att gca ctt cca tca 48
Met Gly Lys Leu Leu Ile Leu Gly Ser Ala Ile Ala Leu Pro Ser
1 5 10 15

ttt gct gct gct ggt ggc gat cta gat ata agt gat act gtt ggt gtt 96
Phe Ala Ala Gly Gly Asp Leu Asp Ile Ser Asp Thr Val Gly Val
20 25 30

tca ttc tgg ctg gtt aca gct ggt atg tta gcg gca act gta ttc ttt 144
Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Thr Val Phe
35 40 45

ttt gta gaa aga gac caa gtc agc gct aag tgg aaa act tca ctt act 192
Phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Thr
50 55 60

gta tct ggt tta att act ggt ata gct ttt tgg cat tat ctc tac atg 240
Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met
65 70 75 80

aga ggt gtt tgg ata gat act ggt gat aca cca gta ttt aga tat 288
Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr
85 90 95

att gat tgg tta tta act gtt cca tta caa atg gtt gag ttc tat cta 336
Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Met Val Glu Phe Tyr Leu
100 105 110

att ctt gct gct tgt aca agt gtt gct gct tca tta ttt aag aag ctt 384
Ile Leu Ala Ala Cys Thr Ser Val Ala Ala Ser Leu Phe Lys Lys Leu
```

115	120	125	
cta gct ggt tca tta gta atg tta ggt gct gga ttt gca ggc gaa gct			432
Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Phe Ala Gly Glu Ala			
130	135	140	
ggt tta gct cct gta tta cct gct ttc att att ggt atg gct gga tgg			480
Gly Leu Ala Pro Val Leu Pro Ala Phe Ile Ile Gly Met Ala Gly Trp			
145	150	155	160
tta tac atg att tat gag cta cat atg ggt gaa ggt aag gct gct gta			528
Leu Tyr Met Ile Tyr Glu Leu His Met Gly Glu Gly Lys Ala Ala Val			
165	170	175	
agt act gca agt cct gct gtt aac tct gca tac aac gca atg atg aag			576
Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Lys			
180	185	190	
att att gtt att gga tgg gca att tat cct gct gga tat gct gct ggt			624
Ile Ile Val Ile Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly			
195	200	205	
tac cta atg agt ggt gac ggt gta tac gct tca aac tta aac ctt ata			672
Tyr Leu Met Ser Gly Asp Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile			
210	215	220	
tat aac ctt gct gac ttt gtt aac aag att cta ttt ggt ttg atc att			720
Tyr Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile			
225	230	235	240
tgg aat gtt gct gtt aaa gaa tct tct aat gct			753
Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala			
245	250		

09847513.000001

<210> 19  
 <211> 251  
 <212> PRT  
 <213> Naturally occurring gamma proteobacterium  
  
 <400> 19  
  
 Met Gly Lys Leu Leu Ile Leu Gly Ser Ala Ile Ala Leu Pro Ser  
 1 5 10 15  
  
 Phe Ala Ala Gly Gly Asp Leu Asp Ile Ser Asp Thr Val Gly Val  
 20 25 30  
  
 Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Ala Thr Val Phe Phe  
 35 40 45  
  
 Phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
 50 55 60  
  
 Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met  
 65 70 75 80  
  
 Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr  
 85 90 95  
  
 Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Met Val Glu Phe Tyr Leu  
 100 105 110

09/847,513,000001

Ile Leu Ala Ala Cys Thr Ser Val Ala Ala Ser Leu Phe Lys Lys Leu  
115 120 125

Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Phe Ala Gly Glu Ala  
130 135 140

Gly Leu Ala Pro Val Leu Pro Ala Phe Ile Ile Gly Met Ala Gly Trp  
145 150 155 160

Leu Tyr Met Ile Tyr Glu Leu His Met Gly Glu Gly Lys Ala Ala Val  
165 170 175

Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Lys  
180 185 190

Ile Ile Val Ile Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly  
195 200 205

Tyr Leu Met Ser Gly Asp Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile  
210 215 220

Tyr Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile  
225 230 235 240

Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala  
245 250

09847513.000001

<210> 20  
 <211> 753  
 <212> DNA  
 <213> Naturally occurring gamma proteobacterium  
  
 <220>  
 <221> CDS  
 <222> (1)..(753)  
 <223> Proteorhodopsin variant from pcr clone HOT75m4; GenBank #AF349981  
  
 <400> 20 48  
 atg ggt aaa tta tta ctg ata tta ggt agt gct att gca ctt cca tca  
 Met Gly Lys Leu Leu Ile Leu Gly Ser Ala Ile Ala Leu Pro Ser  
 1 5 10 15  
  
 ttt gct gct gct ggt ggc gat cta gat ata agt gat act gtt ggt gtt 96  
 Phe Ala Ala Ala Gly Gly Asp Leu Asp Ile Ser Asp Thr Val Gly Val  
 20 25 30  
  
 tca ttc tgg ctg gtt aca gct ggt atg tta gcg gca act gtg ttc ttt 144  
 Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Ala Thr Val Phe Phe  
 35 40 45  
  
 ttt gta gaa aga gac caa gtc agc gct aag gct aaa act tca ctt act 192  
 Phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
 50 55 60  
  
 gta tct ggt tta att act ggt ata gct ttt tgg cat tat ctc tat atg 240  
 Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met  
 65 70 75 80  
  
 aga ggt gtt tgg ata gac act ggt gat acc cca aca gta ttc aga tat 288  
 Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr

09847513.0000001



	85	90	95	
att gat tgg tta tta act gtt cca tta caa gtg gtt gag ttc tat cta				336
Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Val Glu Phe Tyr Leu				
	100	105	110	
att ctt gct gct tgt aca agt gtt gct gct tta ttt aag aag ctt				384
Ile Leu Ala Ala Cys Thr Ser Val Ala Ala Ser Leu Phe Lys Lys Leu				
	115	120	125	
cta gct ggt tca tta gta atg tta ggt gct gga ttt gca ggc gaa gct				432
Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Phe Ala Gly Glu Ala				
	130	135	140	
gga tta gct cct gta tta cct gct ttc att att ggt atg gct gga tgg				480
Gly Leu Ala Pro Val Leu Pro Ala Phe Ile Ile Gly Met Ala Gly Trp				
	145	150	155	
tta tac atg att tat gag cta tat atg ggt gaa ggt aag gct gct gta				528
Leu Tyr Met Ile Tyr Glu Leu Tyr Met Gly Glu Lys Ala Ala Val				
	165	170	175	
agt act gca agt cct gct gtt aac tct gca tac aac gca atg atg atg				576
Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Met				
	180	185	190	
att att gtt gtt gga tgg gca att tat cct gct gga tat gct gct ggt				624
Ile Ile Val Val Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly				
	195	200	205	
tac cta atg ggt ggc gaa ggt gta tac gct tca aac tta aac ctt ata				672
Tyr Leu Met Gly Gly Glu Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile				
	210	215	220	
tat aac ctt gct gac ttt gtt aac aag att cta ttt ggt ttg atc att				720

Tyr Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile  
225 230 235 240

tgg aat gtt gct gtt aaa gaa tct tct aat gct 753

Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala  
245 250

<210> 21

<211> 251

<212> PRT

<213> Naturally occurring gamma proteobacterium

<400> 21

Met Gly Lys Leu Leu Leu Ile Leu Gly Ser Ala Ile Ala Leu Pro Ser  
1 5 10 15

Phe Ala Ala Ala Gly Gly Asp Leu Asp Ile Ser Asp Thr Val Gly Val  
20 25 30

Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Ala Thr Val Phe Phe  
35 40 45

Phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
50 55 60

Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met  
65 70 75 80

Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr  
 85 90 95  
 Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Val Val Glu Phe Tyr Leu  
 100 105 110  
 Ile Leu Ala Ala Cys Thr Ser Val Ala Ala Ser Leu Phe Lys Lys Leu  
 115 120 125  
 Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Phe Ala Gly Glu Ala  
 130 135 140  
 Gly Leu Ala Pro Val Leu Pro Ala Phe Ile Ile Gly Met Ala Gly Trp  
 145 150 155 160  
 Leu Tyr Met Ile Tyr Glu Leu Tyr Met Gly Glu Gly Lys Ala Ala Val  
 165 170 175  
 Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Met  
 180 185 190  
 Ile Ile Val Val Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly  
 195 200 205  
 Tyr Leu Met Gly Gly Glu Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile  
 210 215 220

Tyr Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile  
 225 230 235 240

Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala  
 245 250

<210> 22  
 <211> 753  
 <212> DNA  
 <213> Naturally occurring gamma proteobacterium

<220>  
 <221> CDS  
 <222> (1)..(753)  
 <223> Proteorhodopsin variant from pcr clone HOT75m8: GenBank#AF349982

<400> 22 48  
 atg ggt aaa tta tta ctg ata tta ggt agt gct att gca ctt cca tca  
 Met Gly Lys Leu Leu Ile Leu Gly Ser Ala Ile Ala Leu Pro Ser  
 1 5 10 15

ttt gct gct gct ggt ggc gat cta gat ata agt gat act gtt ggt gtt  
 Phe Ala Ala Ala Gly Asp Leu Asp Ile Ser Asp Thr Val Gly Val  
 20 25 30 96

tca ttc tgg ctg gtt aca gct ggt atg tta gcg gca act gtg ttc ttt  
 Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Ala Thr Val Phe  
 35 40 45 144

ttt gta gaa aga gac caa gtc agc gct aag tgg aaa act tca ctt act 192

09847513.0000001

Phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
 50 55 60  
 gta tct ggt tta att act ggt ata gct ttt tgg cat tat ctc tat atg 240  
 Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met  
 65 70 75 80  
 aga ggt gtt tgg ata gac act ggt gat acc cca aca gta ttc aga tat 288  
 Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr  
 85 90 95  
 att gat tgg tta tta act gtt cca tta caa atg gtt gag ttc tat cta 336  
 Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Met Val Glu Phe Tyr Leu  
 100 105 110  
 att ctt gct gct tgt aca aat gtt gct gct tca tta ttt aag aag ctt 384  
 Ile Leu Ala Ala Cys Thr Asn Val Ala Ala Ser Leu Phe Lys Lys Leu  
 115 120 125  
 cta gct ggt tca tta gta atg tta ggt gct gct gga ttt gca ggc gaa gct 432  
 Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Phe Ala Gly Glu Ala  
 130 135 140  
 gga ttg gct cct gta tgg cct gct ttc att att ggt atg gct gga tgg 480  
 Gly Leu Ala Pro Val Trp Pro Ala Phe Ile Ile Gly Met Ala Gly Trp  
 145 150 155 160  
 tta tac atg att tat gag cta tat atg ggt gaa ggt aag gct gct gta 528  
 Leu Tyr Met Ile Tyr Glu Leu Tyr Met Gly Glu Lys Ala Ala Val  
 165 170 175  
 agt act gca agt cct gct gtt aac tct gca tac aac gca atg atg gtg 576  
 Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Val  
 180 185 190

att att gtt gtt gga tgg gca att tat cct gct gga tat gct gct ggt 624  
 ile ile val val gly trp ala ile tyr pro ala gly tyr ala ala gly  
 195 200 205

tac cta atg ggt ggc gaa ggt gta tac gct tca aac tta aac ctt ata 672  
 tyr leu met gly gly glu gly val tyr ala ser asn leu asn leu ile  
 210 215 220

tat aac ctt gcc gac ctt gtt aac aag att cta ttt ggt ttg atc att 720  
 tyr asn leu ala asp leu val asn lys ile leu phe gly leu ile ile  
 225 230 235 240

tgg aat gtt gct gtt aaa gaa tct tct aat gct 753  
 trp asn val ala val lys glu ser ser asn ala  
 245 250

<210> 23

<211> 251

<212> PRT

<213> Naturally occurring gamma proteobacterium

<400> 23

Met Gly Lys Leu Leu Leu Ile Leu Gly Ser Ala Ile Ala Leu Pro Ser  
 1 5 10 15

Phe Ala Ala Ala Gly Gly Asp Leu Asp Ile Ser Asp Thr Val Gly Val  
 20 25 30

Ser Phe Trp Trp Leu Val Thr Ala Gly Met Leu Ala Ala Thr Val Phe Phe  
 35 40 45

Phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
50 55 60  
Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met  
65 70 75 80  
Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr  
85 90 95  
Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Met Val Glu Phe Tyr Leu  
100 105 110  
Ile Leu Ala Ala Cys Thr Asn Val Ala Ala Ser Leu Phe Lys Lys Leu  
115 120 125  
Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Phe Ala Gly Glu Ala  
130 135 140  
Gly Leu Ala Pro Val Trp Pro Ala Phe Ile Ile Gly Met Ala Gly Trp  
145 150 155 160  
Leu Tyr Met Ile Tyr Glu Leu Tyr Met Gly Glu Lys Ala Ala Val  
165 170 175  
Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Val  
180 185 190

09847513.00000001

Ile Ile Val Val Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly  
195 200 205

Tyr Leu Met Gly Gly Glu Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile  
210 215 220

Tyr Asn Leu Ala Asp Leu Val Asn Lys Ile Leu Phe Gly Leu Ile Ile  
225 230 235 240

Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala  
245 250

<210> 24

<211> 750

<212> DNA

<213> Naturally occurring gamma proteobacterium

<220>

<221> CDS

<222> (1)..(750)

<223> Proteorhodopsin variant from pcr clone MB0m1: GenBank#AF349983

<400> 24

atg ggt aaa tta tta ctg ata tta ggt agt gtt att gca ctt cct aca 48

Met Gly Lys Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr  
1 5 10 15

ttt gct gca ggt ggt ggt gac ctt gat gct agt gat tac act ggt gtt 96



Phe Ala Ala Gly Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val  
 20 25 30  
 144  
 tct ttt tgg tta gtt act gct gct cta tta gca tct act gta ttt ttc  
 Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe  
 35 40 45  
 192  
 ttt gtt gaa aga gat aga gtt tct gca aaa tgg aaa tca tta act  
 Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
 50 55 60  
 240  
 gta tct ggt ctt gtt act ggt att gct ttc tgg cat tac atg tac atg  
 Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met  
 65 70 75 80  
 288  
 aga ggg gta tgg att gag act ggt gat tcg cca act gta ttt aga tac  
 Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr  
 85 90 95  
 336  
 att gat tgg tta cta aca gtt cct cta ttg ata tgt gaa ttc tac tta  
 Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu  
 100 105 110  
 384  
 att ctt gct gct gca aca aat gtt gct gct ggc ctg ttt aag aaa tta  
 Ile Leu Ala Ala Thr Asn Val Ala Ala Gly Leu Phe Lys Lys Leu  
 115 120 125  
 432  
 ttg gtt ggt tct ctt gtt atg ctt gtt ggt tac atg ggt gag gca  
 Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala  
 130 135 140  
 480  
 gga att atg aac gct tgg cct gca ttc att att ggg tgt tta gct tgg  
 Gly Ile Met Asn Ala Trp Pro Ala Phe Ile Ile Gly Cys Leu Ala Trp  
 145 150 155 160

147/235  
 147/235

528 gta tac atg att tat gaa cta tat gct gga gaa gga aaa tct gca tgt  
 Val Tyr Met Ile Tyr Glu Leu Tyr Ala Gly Glu Gly Lys Ser Ala Cys  
 165 170 175  
  
 576 aat act gca agt cct tcg gtt caa tca gct tac aac aca atg atg gct  
 Asn Thr Ala Ser Pro Ser Val Gln Ser Ala Tyr Asn Thr Met Met Ala  
 180 185 190  
  
 624 atc ata gtc ttc ggt tgg gca att tat cct gta ggt tat ttc aca ggt  
 Ile Ile Val Phe Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly  
 195 200 205  
  
 672 tac cta atg ggt gac ggt gga tca gct ctt aac tta aac ctt att tat  
 Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Ile Tyr  
 210 215 220  
  
 720 aac ctt gct gac ttt gtt aac aag att cta ttt ggt tta att ata tgg  
 Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile Trp  
 225 230 235 240  
  
 750 aat gtt gct gtt aaa gaa tct tct aat gct  
 Asn Val Ala Val Lys Lys Glu Ser Ser Asn Ala  
 245 250  
  
 <210> 25  
 <211> 250  
 <212> PRT  
 <213> Naturally occurring gamma proteobacterium  
  
 <400> 25  
 Met Gly Lys Leu Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr  
 1 5 10 15

Phe Ala Ala Gly Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val  
 20 25 30  
 Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe  
 35 40 45  
 Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
 50 55 60  
 Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met  
 65 70 75 80  
 Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr  
 85 90 95  
 Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu  
 100 105 110  
 Ile Leu Ala Ala Ala Thr Asn Val Ala Ala Gly Leu Phe Lys Lys Leu  
 115 120 125  
 Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala  
 130 135 140  
 Gly Ile Met Asn Ala Trp Pro Ala Phe Ile Ile Gly Cys Leu Ala Trp  
 145 150 155 160

Val Tyr Met Ile Tyr Glu Leu Tyr Ala Gly Glu Lys Ser Ala Cys  
165 170 175

Asn Thr Ala Ser Pro Ser Val Gln Ser Ala Tyr Asn Thr Met Met Ala  
180 185 190

Ile Ile Val Phe Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly  
195 200 205

Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Ile Tyr  
210 215 220

Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile Trp  
225 230 235 240

Asn Val Ala Val Lys Glu Ser Ser Asn Ala  
245 250

<210> 26  
<211> 750  
<212> DNA  
<213> Naturally occurring gamma proteobacterium

<220>  
<221> CDS  
<222> (1)..(750)  
<223> Proteorhodopsin variant from pcr clone MB0m2

<400> 26  
 atg ggt aaa tta tta ctg ata tta ggt agt gtt att gca ctt cct aca  
 Met Gly Lys Leu Leu 5  
 1  
 ttt gct gca ggt ggt ggt gac ctt gat gct agt gat tac act ggt gtt  
 Phe Ala Ala Gly Gly Asp Leu 20  
 96  
 tct ttt tgg tta gtt act gct gct tta tta gca tct act gta ttt ttc  
 Ser Phe Trp Leu Val Thr Ala Ala Leu 35  
 144  
 ttt gtt gaa aga gat aga gtt tct gca aaa tgg aaa aca tca tta act  
 Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp 50  
 192  
 gta tct ggt ctt gtt act ggt att gct ttc tgg cat tac atg tac atg  
 Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr 65  
 240  
 aga ggg gta tgg att gaa act ggt gat tcg cca act gta ttt aga tac  
 Arg Gly Val Trp 85  
 288  
 att gat tgg tta cta aca gtt cct cta tta ata tgt gaa ttc tac tta  
 Ile Asp Trp Leu Leu Thr Val Pro Leu 100  
 336  
 att ctt gct gct act aat gtt gct gct ggc ctg ttt aag aaa tta  
 Ile Leu Ala Ala Ala Thr Asn Val Ala Ala Gly Leu Phe 115  
 384



<212> PRT  
<213> Naturally occurring gamma proteobacterium

<400> 27

Met Gly Lys Leu Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr  
1 5 10 15

Phe Ala Ala Gly Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val  
20 25 30

Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe  
35 40 45

Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
50 55 60

Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met  
65 70 75 80

Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr  
85 90 95

Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu  
100 105 110

Ile Leu Ala Ala Ala Thr Asn Val Ala Ala Gly Leu Phe Lys Lys Leu  
115 120 125

Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala  
130 135 140

Gly Ile Met Asn Ala Trp Gly Ala Phe Val Ile Gly Cys Leu Ala Trp  
145 150 155 160

Val Tyr Met Ile Tyr Glu Leu Trp Leu Gly Glu Lys Ala Ala Cys  
165 170 175

Asn Thr Ala Ser Pro Ala Val Gln Ser Ala Tyr Asn Thr Met Met Met  
180 185 190

Ile Ile Ile Phe Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly  
195 200 205

Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Ile Tyr  
210 215 220

Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile Trp  
225 230 235 240

Asn Val Ala Val Lys Glu Ser Ser Asn Ala  
245 250

<210> 28

154/235



<211> 750  
 <212> DNA  
 <213> Naturally occurring gamma proteobacterium  
  
 <220>  
 <221> CDS  
 <222> (1)..(750)  
 <223> Proteorhodopsin variant from pcr clone MB20m2; GenBank #AF349985  
  
 <400> 28 48  
 atg ggt aaa tta tta ctg ata tta ggt agt gtt att gca ctt cct aca  
 Met Gly Lys Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr  
 1 5 10 15  
  
 ttt gct gca ggt ggt ggt gac ctt gat gct agt gat tac act ggt gtt  
 Phe Ala Ala Gly Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val  
 20 25 30  
  
 tct ttt tgg tta gtt act gct gct tta tta gca tct act gta ttt ttc  
 Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe  
 35 40 45  
  
 ttt gtt gaa aga gat aga gtt tct gca aaa tgg aaa aca tca tta act  
 Phe Val Glu Arg Arg Asp Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
 50 55 60  
  
 gta tct ggt ctt gtt act ggt att gct ttc ttc tgg cat tac atg tac atg  
 Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met  
 65 70 75 80  
  
 aga ggg gta tgg att gaa act ggt gat tcg cca act gta ttt aga tac  
 Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr  
 85 90 95  
 288

att gat tgg tta cta aca gtt cct cta tta ata tgt gaa ttc tac tta  
 Ile Asp Trp Leu Leu Thr Val Pro Leu Ile Cys Glu Phe Tyr Leu  
 100 105 110 336  
  
 att ctt gct gct gca act aat gtt gct gct ggc ctg ttt aag aaa tta  
 Ile Leu Ala Ala Thr Asn Val Ala Ala Gly Leu Phe Lys Lys Leu  
 115 120 125 384  
  
 ttg gtt ggt tct ctt gtt atg ctt gtg ttt ggt tac atg ggt gag gca  
 Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala  
 130 135 140 432  
  
 gga att atg aac gct tgg ggt gca ttc gtt att ggg tgt tta gct tgg  
 Gly Ile Met Asn Ala Trp Gly Ala Phe Val Ile Gly Cys Leu Ala Trp  
 145 150 155 160 480  
  
 gta tac atg att tat gaa cta tgg gct gga gaa ggc aag gct gca tgt  
 Val Tyr Met Ile Tyr Glu Leu Trp Ala Gly Glu Lys Ala Ala Cys  
 165 170 175 528  
  
 aat act gca agt cct gct gtg caa tca gct tac aac aca atg atg tat  
 Asn Thr Ala Ser Pro Ala Val Gln Ser Ala Tyr Asn Thr Met Met Tyr  
 180 185 190 576  
  
 ata atc atc ttt ggt tgg gca att tat cct gta ggt tat ttc aca ggt  
 Ile Ile Ile Phe Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly  
 195 200 205 624  
  
 tac cta atg ggt gac ggt gga tca gct ctt aac tta aac ctt atc tat  
 Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Asn Ile Tyr  
 210 215 220 672  
  
 aac ctt gct gac ttt gtt aac aag att cta ttt ggt tta att ata tgg  
 Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile Trp  
 720

09/847,513 DeLong et al.

225	230	235	240	750
aat gtt gct gtt aaa gaa tct tct aat gct				
Asn Val Ala Val Lys Glu Ser Ser Asn Ala				
245	250			
<210> 29				
<211> 250				
<212> PRT				
<213> Naturally occurring gamma proteobacterium				
<400> 29				
Met Gly Lys Leu Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr				
1	5	10	15	
Phe Ala Ala Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val				
20	25	30		
Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe				
35	40	45		
Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr				
50	55	60		
Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met				
65	70	75	80	
Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr				

090347Z FEB 68 080800Z

Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu	85	90	95
100	105	110	
Ile Leu Ala Ala Thr Asn Val Ala Ala Gly Leu Phe Lys Lys Leu	115	120	125
130	135	140	
Gly Ile Met Asn Ala Trp Gly Ala Phe Val Ile Gly Cys Leu Ala Trp	145	150	155
160	165	170	175
Val Tyr Met Ile Tyr Glu Leu Trp Ala Gly Glu Lys Ala Ala Cys	180	185	190
Asn Thr Ala Ser Pro Ala Val Gln Ser Ala Tyr Asn Thr Met Met Tyr	195	200	205
Ile Ile Ile Phe Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly	210	215	220
Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Ile Tyr	225	230	235

09847513.0000001

Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile Trp  
 225 230 235 240

Asn Val Ala Val Lys Glu Ser Ser Asn Ala  
 245 250

<210> 30  
 <211> 750  
 <212> DNA  
 <213> Naturally occurring gamma proteobacterium

<220>  
 <221> CDS  
 <222> (1)..(750)  
 <223> Proteorhodopsin variant from pcr clone MB20m5; GenBank#AF349986

<400> 30 48  
 atg ggt aaa tta tta ctg ata tta ggt agt gtt att gca ctt cct aca  
 Met Gly Lys Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr  
 1 5 10 15

ttt gct gca ggt ggc ggt gac ctt gat gct agt gat tac act ggt gtt  
 Phe Ala Ala Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val  
 20 25 30 96

tct ttt tgg tta gtt aca gct gct cta tta gca tct act gta ttt ttc  
 Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe  
 35 40 45 144

ttt gtt gaa aga gat aga gtt tct gca aaa tgg aaa aca tta act  
 Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
 50 55 60 192

09847513.0000001



195	200	205	672
tac cta atg ggt gac ggt ggg tca gct ctt aac tta aac ctt att tat			
Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Ile Tyr			
210	215	220	720
aac ctt gct gac ttt gtt aac aag att cta ctt ggt tta att ata tgg			
Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Leu Gly Leu Ile Ile Trp			
225	230	235	750
aat gtt gct gtt aaa gaa tct tct aat gct			
Asn Val Ala Val Lys Glu Ser Ser Asn Ala			
	245	250	

31  
 <210>  
 250  
 <211>  
 PRT  
 <212>  
 <213> Naturally occurring gamma proteobacterium  
  
 31  
 <400>  
 Met Gly Lys Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr  
 5 10 15  
 1  
 Phe Ala Ala Gly Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val  
 20 25 30  
 Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe  
 35 40 45  
 Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr

Corrected Sequence Listing (August 4<sup>th</sup>, 2001)

09/847,513 DeLong et al.





Ile Ile Val Phe Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly  
195 200 205

Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Ile Tyr  
210 215 220

Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Leu Gly Ile Ile Trp  
225 230 235 240

Asn Val Ala Val Lys Glu Ser Ser Asn Ala  
245 250

<210> 32

<211> 750

<212> DNA

<213> Naturally occurring gamma proteobacterium

<220>

<221> CDS

<222> (1)..(750)

<223> Proteorhodopsin variant from pcr clone MB20m12; GenBank #AF349987

<400> 32

atg ggt aaa tta tta ctg ata tta ggt agt gtt att gca ctt cct aca  
Met Gly Lys Leu Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr  
1 5 10 15 48

ttt gct gca ggt ggt ggt gac ctt gat gct agt gat tac act ggt gtt  
Phe Ala Ala Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val  
20 25 30 96

tct ttt tgg tta gtt act gct gct tta tta gca tct act gta ttt ttc 144  
Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe  
35 40 45  
  
ttt gtt gaa aga gat aga gtt tct gca aaa tgg aaa aca tta act 192  
Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
50 55 60  
  
gta tct ggt ctt gtt act ggt att gct ttc tgg cat tac atg tac atg 240  
Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met  
65 70 75 80  
  
aga ggg gta tgg att gaa act ggt gat tcg cca act gta ttt aga tac 288  
Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr  
85 90 95  
  
att gat tgg tta cta aca gtt cct cta tta ata tgt gaa ttc tac tta 336  
Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu  
100 105 110  
  
att ctt gct gct gca gct aat gtt gct gga tca tta ttt aag aaa tta 384  
Ile Leu Ala Ala Ala Asn Val Ala Gly Ser Leu Phe Lys Lys Leu  
115 120 125  
  
cta gtt ggt tct ctt gtt atg ctt gtg ttt ggt tac atg ggt gaa gca 432  
Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala  
130 135 140  
  
gga atc atg gct gca tgg cct gca ttc att att ggg tgt tta gct tgg 480  
Gly Ile Met Ala Ala Trp Pro Ala Phe Ile Ile Gly Cys Leu Ala Trp  
145 150 155 160  
  
gta tac atg att tat gaa tta tgg gct gga gaa gaa tct gca tgt 528  
Val Tyr Met Ile Tyr Glu Leu Trp Ala Gly Glu Gly Lys Ser Ala Cys

165	170	175	576
aat act gca agt cct gct gtg caa tca gcc tac aac aca atg atg tat			
Asn Thr Ala Ser Pro Ala Val Gln Ser Ala Tyr Asn Thr Met Met Tyr			
180	185	190	
att atc atc ttt ggt tgg gcg att tat cct gta ggt tat ttc aca ggt			
Ile Ile Ile Phe Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly			
195	200	205	624
tac ttg atg ggt gac ggt gga tca gct ctt aac tta aac ctt atc tat			
Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Ile Tyr			
210	215	220	672
aac ctt gct gac ttt gtt aac aag att cta ttt ggt tta att ata tgg			
Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile Trp			
225	230	235	720
aat gtt gct gtt aaa gaa tct tct aat gct			
Asn Val Ala Val Lys Glu Ser Ser Asn Ala			
245	250		750
<210> 33			
<211> 250			
<212> PRT			
<213> Naturally occurring gamma proteobacterium			
<400> 33			
Met Gly Lys Leu Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr			
1	5	10	15
Phe Ala Ala Gly Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val			

09/847,513 DeLong et al.

Corrected Sequence Listing (August 4<sup>th</sup>, 2001)

165/235

**SECRET**

	20	25	30
Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe			
35	40	45	
Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr			
50	55	60	
Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met			
65	70	75	80
Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr			
85	90	95	
Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu			
100	105	110	
Ile Leu Ala Ala Ala Asn Val Ala Gly Ser Leu Phe Lys Lys Leu			
115	120	125	
Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala			
130	135	140	
Gly Ile Met Ala Ala Trp Pro Ala Phe Ile Ile Gly Cys Leu Ala Trp			
145	150	155	160

09/847,513 . 0000001

Val Tyr Met Ile Tyr Glu Leu Trp Ala Gly Glu Lys Ser Ala Cys  
165 170 175

Asn Thr Ala Ser Pro Ala Val Gln Ser Ala Tyr Asn Thr Met Met Tyr  
180 185 190

Ile Ile Ile Phe Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly  
195 200 205

Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Ile Tyr  
210 215 220

Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile Trp  
225 230 235 240

Asn Val Ala Val Lys Glu Ser Ser Asn Ala  
245 250

<210> 34

<211> 750

<212> DNA

<213> Naturally occurring gamma proteobacterium

<220>

<221> CDS

<222> (1)..(750)

<223> Proteorhodopsin variant from pcr clone MB40ml; GenBank #AF349988

<400> 34  
 atg ggt aaa tta tta ctg ata ata ggt agt gtt att gca ctt cct aca 48  
 Met Gly Lys Leu Leu Ile Ile Gly Ser Val Ile Ala Leu Pro Thr  
 1 5 10 15  
 ttt gct gca ggt ggc ggt gac ctt gat gct agt gat tac act ggt gtt 96  
 Phe Ala Ala Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val  
 20 25 30  
 tct ttt tgg tta gtt aca gct gct cta tta gca tct act gta ttt ttc 144  
 Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe  
 35 40 45  
 ttt gtt gaa aga gat aga gtt tct gca aaa tgg aaa aca tca tta act 192  
 Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
 50 55 60  
 gta tct ggt ctt gtt act ggt att gct ttc tgg cat tac atg tac atg 240  
 Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met  
 65 70 75 80  
 aga gga gta tgg att gaa act ggt gat tcg cca act gta ttt aga tac 288  
 Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr  
 85 90 95  
 att gat tgg tta cta aca gtt cct tta tta ata tgt gaa ttc tac tta 336  
 Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu  
 100 105 110  
 att ctt gct gct gca act aat gtt gcc ggc tca tta ttt aag aaa ctt 384  
 Ile Leu Ala Ala Ala Thr Asn Val Ala Gly Ser Leu Phe Lys Lys Leu  
 115 120 125  
 cta gtt ggt tct ctt atg gtt atg ctt gtg ttt ggt tac atg ggt gaa gca 432

Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala  
 130 135  
 gga att atg gca gct tgg cct gca ttc att att ggg tgt tta gct tgg 480  
 Gly Ile Met Ala Ala Trp Pro Ala Phe Ile Ile Gly Cys Leu Ala Trp  
 145 150 155 160  
 gta tat atg att tat gaa cta tat gct gga gaa gga aaa tct gca tgt 528  
 Val Tyr Met Ile Tyr Glu Leu Tyr Ala Gly Glu Gly Lys Ser Ala Cys  
 165 170 175  
 aat aca gca agt cct gct gtg caa tca gct tac aac aca atg atg tat 576  
 Asn Thr Ala Ser Pro Ala Val Gln Ser Ala Tyr Asn Thr Met Met Tyr  
 180 185 190  
 att atc gtc ttt ggt tgg gcg att tat cct gta ggt tat ttc aca ggt 624  
 Ile Ile Val Phe Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly  
 195 200 205  
 tac ctg atg ggt gac ggt gga tca gct ctt aac tta aac ctt atc tat 672  
 Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Asn Ile Tyr  
 210 215 220  
 aac ctt gct gac ttt gtt aac aag att cta ttt ggt tta att ata tgg 720  
 Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile Trp  
 225 230 235 240  
 aat gtt gct gtt aaa gaa tct tct aat gct 750  
 Asn Val Ala Val Lys Glu Ser Ser Asn Ala  
 245 250

<210> 35  
 <211> 250  
 <212> PRT

<213> Naturally occurring gamma proteobacterium

<400> 35

Met Gly Lys Leu Leu Ile Ile Gly Ser Val Ile Ala Leu Pro Thr  
1 5 10 15

Phe Ala Ala Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val  
20 25 30

Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe  
35 40 45

Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
50 55 60

Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met  
65 70 75 80

Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr  
85 90 95

Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu  
100 105 110

Ile Leu Ala Ala Ala Thr Asn Val Ala Gly Ser Leu Phe Lys Lys Leu  
115 120 125



Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala  
130 135 140

Gly Ile Met Ala Ala Trp Pro Ala Phe Ile Ile Gly Cys Leu Ala Trp  
145 150 155 160

Val Tyr Met Ile Tyr Glu Leu Tyr Ala Gly Glu Lys Ser Ala Cys  
165 170 175

Asn Thr Ala Ser Pro Ala Val Gln Ser Ala Tyr Asn Thr Met Met Tyr  
180 185 190

Ile Ile Val Phe Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly  
195 200 205

Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Ile Tyr  
210 215 220

Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile Trp  
225 230 235 240

Asn Val Ala Val Lys Glu Ser Ser Asn Ala  
245 250

<210> 36

<211> 750

<212> DNA  
 <213> Naturally occurring gamma proteobacterium  
  
 <220>  
 <221> CDS  
 <222> (1)..(750)  
 <223> Proteorhodopsin variant from pcr clone MB40m5;p GenBank #AF349989  
  
 <400> 36 48  
 atg ggt aaa tta tta ctg ata tta ggt agt gtt att gca ctt cct aca  
 Met Gly Lys Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr  
 1 5 10 15  
  
 ttt gct gca ggt ggt ggt gac ctt gat gct agt gat tac act ggt gtt  
 Phe Ala Ala Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val  
 20 25 30 36  
  
 tct ttt tgg tta gtt act gct gct cta tta gca tct act gta ttt ttc  
 Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe  
 35 40 45 51  
  
 ttt gtt gaa aga gat aga gtt tct gca aaa tgg aaa aca tta act  
 Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
 50 55 60 66  
  
 gta tcg ggt ctt gtt act ggt att gct ttc tgg cat tac atg tac atg  
 Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met  
 65 70 75 80 86  
  
 aga ggg gta tgg att gag act ggt gat tcg cca act gta ttt aga tac  
 Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr  
 85 90 95 101  
  
 att gat tgg tta cta aca gtt cct cta ttg ata tgt gaa ttc tac tta  
 106 112 118 124 130 136 142

09847513 DeLong et al.



750

aat gtt gct gtt aaa gaa tct tct aat gct  
Asn Val Ala Val Lys Glu Ser Ser Asn Ala  
245 250

<210> 37

<211> 250

<212> PRT

<213> Naturally occurring gamma proteobacterium

<400> 37

Met Gly Lys Leu Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr  
1 5 10 15

Phe Ala Ala Gly Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val  
20 25 30

Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe  
35 40 45

Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
50 55 60

Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met  
65 70 75 80

Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr  
85 90 95

09847513.10000001

Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu  
 100 105 110  
  
 Ile Leu Ala Ala Thr Asn Val Ala Ala Gly Leu Phe Lys Lys Leu  
 115 120 125  
  
 Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala  
 130 135 140  
  
 Gly Ile Met Asn Ala Trp Gly Ala Phe Val Ile Gly Cys Leu Ala Trp  
 145 150 155 160  
  
 Val Tyr Met Ile Tyr Glu Leu Trp Ala Gly Glu Gly Lys Ala Ala Cys  
 165 170 175  
  
 Asn Thr Ala Ser Pro Ala Val Gln Ser Ala Tyr Asn Thr Met Met Tyr  
 180 185 190  
  
 Ile Ile Ile Phe Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly  
 195 200 205  
  
 Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Asn Leu Ile Tyr  
 210 215 220  
  
 Asn Leu Ala Asp Phe Val Asn Lys Asn Leu Phe Gly Leu Ile Ile Trp  
 225 230 235 240

09847513.00000001

Asn Val Ala Val Lys Glu Ser Ser Asn Ala  
245 250

<210> 38  
<211> 750  
<212> DNA  
<213> Naturally occurring gamma proteobacterium

<220>  
<221> CDS  
<222> (1)..(750)  
<223> Proteorhodopsin variant from pcr clone MB40m12; GenBank # AF34999

<400> 38 48  
atg ggt aaa tta tta cgg ata tta ggt agt gtt att gca ctt cct aca  
Met Gly Lys Leu Leu Arg Ile Leu Gly Ser Val Ile Ala Leu Pro Thr  
1 5 10 15

ttt gct gca ggt ggc ggt gac ctt gat gct agt gat tac act ggt gtt 96  
Phe Ala Ala Gly Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val  
20 25 30

tct ttt tgg tta gtt aca gct gct cta tta gca tct act gta ttt ttc 144  
Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe  
35 40 45

ttt gtt gaa aga gat aga gtt tct gca aaa tgg aaa aca tta act 192  
Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
50 55 60

gta tct ggt ctt gtt act ggt att gct ttc tgg cat tac atg tat atg 240  
 Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met 80  
 65 70 75  
 aga gga gta tgg att gaa act ggt gat tcg cca act gta ttt aga tac 288  
 Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr 95  
 85 90  
 att gat tgg tta cta aca gtt cct tta tta ata tgt gaa ttc tac tta 336  
 Ile Asp Trp Leu Leu Thr Val Pro Leu Ile Cys Glu Phe Tyr Leu 110  
 100 105  
 att ctt gct gct gca act aat gtt gct gga tca tta ttt aag aaa tta 384  
 Ile Leu Ala Ala Thr Asn Val Ala Gly Ser Leu Phe Lys Lys Leu 125  
 115 120  
 cta gtt ggt tct ctt gtt atg ctt gtg ttt ggt tac atg ggt gaa gca 432  
 Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala 140  
 130 135  
 gga atc atg gct gca tgg cct gca ttc att att ggg tgt tta gct tgg 480  
 Gly Ile Met Ala Ala Trp Pro Ala Phe Ile Ile Gly Cys Leu Ala Trp 160  
 145 150 155  
 gta tac atg att tat gaa cta tgg gct gga gaa gga aaa tct gca tgt 528  
 Val Tyr Met Ile Tyr Glu Leu Trp Ala Gly Glu Lys Ser Ala Cys 175  
 165 170  
 aat act gca agt cct gct gtg caa tca gct tac aac aca atg atg tat 576  
 Asn Thr Ala Ser Pro Ala Val Gln Ser Ala Tyr Asn Thr Met Met Tyr 190  
 180 185  
 atc atc atc gtt ggt tgg gcg att tat cct gta ggt tat ttc aca ggt 624  
 Ile Ile Ile Val Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly 205  
 195 200

tac ctg atg ggt gac ggt gga tca gct ctt aac tta aac ctt atc tat 672  
 Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Asn Leu Ile Tyr  
 210 215 220  
 aac ctt gct gac ttt gtt aac aag att cta ttt ggt tta att ata tgg 720  
 Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile Trp  
 225 230 235 240  
 aat gtt gct gtt aaa gaa tct tct aat gct 750  
 Asn Val Ala Val Lys Glu Ser Ser Asn Ala 245 250  
  
 <210> 39  
 <211> 250  
 <212> PRT  
 <213> Naturally occurring gamma proteobacterium  
  
 <400> 39  
 Met Gly Lys Leu Leu Arg Ile Leu Gly Ser Val Ile Ala Leu Pro Thr  
 1 5 10 15  
 Phe Ala Ala Gly Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val  
 20 25 30  
 Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe  
 35 40 45  
 Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
 50 55 60



Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met  
65 70 75 80

Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr  
85 90 95

Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu  
100 105 110

Ile Leu Ala Ala Thr Asn Val Ala Gly Ser Leu Phe Lys Lys Leu  
115 120 125

Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala  
130 135 140

Gly Ile Met Ala Ala Trp Pro Ala Phe Ile Ile Gly Cys Leu Ala Trp  
145 150 155 160

Val Tyr Met Ile Tyr Glu Leu Trp Ala Gly Glu Gly Lys Ser Ala Cys  
165 170 175

Asn Thr Ala Ser Pro Ala Val Gln Ser Ala Tyr Asn Thr Met Met Tyr  
180 185 190

Ile Ile Ile Val Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly

09/847,513 DeLong et al.

195

200

205

Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Asn Leu Ile Tyr  
210 215 220

Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile Trp  
225 230 235 240

Asn Val Ala Val Lys Glu Ser Ser Asn Ala  
245 250

<210> 40

<211> 750

<212> DNA

<213> Naturally occurring gamma proteobacterium

<220>

<221> CDS

<222> (1)..(750)

<223> Proteorhodopsin variant from pcr clone MB100m5; GenBank #AF349991

<400> 40

atg ggt aaa tta tta ctg ata tta ggt agt gtt att gca ctt cct aca  
Met Gly Lys Leu Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr  
1 5 10 15

48

ttt gct gca ggt ggc ggt gac ctt gat gct agt gat tac act ggt gtt  
Phe Ala Ala Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val  
20 25 30

96

09847513.0000001

tct ttt tgg tta gtt aca gct gct cta tta gca tct act gta ttt ttc 144  
 Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe  
 35 40 45  
 ttt gtt gaa aga gat aga gtt tct gca aaa tgg aaa aca tca tta act 192  
 Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
 50 55 60  
 gta tct ggt ctt gtt act ggt att gct ttc tgg cat tac atg tac atg 240  
 Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met  
 65 70 75 80  
 aga gga gta tgg att gaa act ggt gat tcg cca act gta ttt aga tac 288  
 Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr  
 85 90 95  
 att gat tgg tta cta aca gtt cct tta tta ata tgt gaa ttc tac tta 336  
 Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu  
 100 105 110  
 att ctt gct gct gca act aat gtt gcc ggc tca tta ttt aag aaa ctt 384  
 Ile Leu Ala Ala Ala Thr Asn Val Ala Gly Ser Leu Phe Lys Lys Leu  
 115 120 125  
 cta gtt ggt tct ctt gtt atg ctt gtg ttt ggt tac atg ggt gaa gca 432  
 Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala  
 130 135 140  
 gga att atg gca gct tgg cct gca ttc att att ggg tgt tta gct tgg 480  
 Gly Ile Met Ala Ala Trp Pro Ala Phe Ile Ile Gly Cys Leu Ala Trp  
 145 150 155 160  
 gta tac atg att tat gaa cta tat gct gga gaa gga aaa tct gca tgt 528  
 Val Tyr Met Ile Tyr Glu Leu Tyr Ala Gly Glu Gly Lys Ser Ala Cys  
 165 170 175

aat act gca agt cct tcg gtt caa tca gct tac aac aca atg atg gct 576  
 Asn Thr Ala Ser Pro Ser Val Gln Ser Ala Tyr Asn Thr Met Met Ala  
 180 185 190  
 atc ata gtc ttc ggt tgg gca att tat cct gta ggt tat ttc aca ggt 624  
 Ile Ile Val Phe Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly  
 195 200 205  
 tac cta atg ggt gac ggt gga tca gct ctt aac tta aac ctt att tat 672  
 Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Asn Leu Ile Tyr  
 210 215 220  
 aac ctt gct gac ttt gtt aac aag att cta ttt ggt tta att ata tgg 720  
 Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile Trp  
 225 230 235 240  
 aat gtt gct gtt aaa gaa tct tct aat gct 750  
 Asn Val Ala Val Lys Glu Ser Ser Asn Ala  
 245 250

<210> 41  
 <211> 250  
 <212> PRT  
 <213> Naturally occurring gamma proteobacterium

<400> 41  
 Met Gly Lys Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr  
 1 5 10 15

Phe Ala Ala Gly Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val  
 20 25 30

Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe  
 35 40 45  
 Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
 50 55 60  
 Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met  
 65 70 75 80  
 Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr  
 85 90 95  
 Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu  
 100 105 110  
 Ile Leu Ala Ala Thr Asn Val Ala Gly Ser Leu Phe Lys Lys Leu  
 115 120 125  
 Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala  
 130 135 140  
 Gly Ile Met Ala Ala Trp Pro Ala Phe Ile Ile Gly Cys Leu Ala Trp  
 145 150 155 160  
 Val Tyr Met Ile Tyr Glu Leu Tyr Ala Gly Glu Lys Ser Ala Cys

09847513.0000001



atg ggt aaa tta tta ctg ata tta ggt agt ggt att gca ctt cct aca 48  
 Met Gly Lys Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr 15  
 1  
 ttt gct gca ggt ggt ggt gac ctt gat gct agt gat tac act ggt gtt 96  
 Phe Ala Ala Gly Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val 30  
 20 25  
 tct ttt tgg tta gtt act gct gct tta tta gca tct act gta ttt ttc 144  
 Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe 45  
 35 40  
 ttt gtt gaa aga gat aga gtt tct gca aaa tgg aaa aca tca tta act 192  
 Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr 60  
 50 55  
 gta tct ggt ctt gtt act ggt att gct ttc tgg cat tac atg tac atg 240  
 Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met 80  
 65 70  
 aga ggg gta tgg att gaa act ggt gat tcg cca act gta ttt aga tac 288  
 Arg Gly Val Trp Ile Glu Thr Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr 90  
 85 95  
 att gat tgg tta cta aca gtt cct cta tta ata tgt gaa ttc tac tta 336  
 Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu 110  
 100 105  
 att ctt gct gct gct act aat gtt gcc ggc tca tta ttt aag aaa ctt 384  
 Ile Leu Ala Ala Ala Thr Asn Val Ala Gly Ser Leu Phe Lys Lys Leu 120  
 115 125  
 cta gtt ggt tct ctt gtt atg ctt gtt ttt ggt tac atg ggt gaa gca 432  
 Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala 140  
 130 135

gga att atg gca gct tgg cct gca ttc att att ggg tgt tta gct tgg 480  
 Gly Ile Met Ala Ala Trp Pro Ala Phe Ile Ile Gly Cys Leu Ala Trp 160  
 145 150 155  
 gta tac atg att tat gaa cta tat gct gga gaa gga aaa tct gca tgt 528  
 Val Tyr Met Ile Tyr Glu Leu Tyr Ala Gly Glu Gly Lys Ser Ala Cys 175  
 165 170  
 aat act gca agt cct tcg gtt caa tca gct tac aac aca atg atg gct 576  
 Asn Thr Ala Ser Pro Ser Val Gln Ser Ala Tyr Asn Thr Met Met Ala 190  
 180 185  
 atc ata gtc ttc ggt tgg gca att tat cct gta ggt tat ttc aca ggt 624  
 Ile Ile Val Phe Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly 205  
 195 200  
 tac cta atg ggt gac ggt gga tca gct ctt aac tta aac ctt att tat 672  
 Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Ile Tyr 215  
 210  
 aac ctt gct gac ttt gtt aac aag att cta ttt ggt tta att ata tgg 720  
 Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile Trp 240  
 225 230 235  
 aat gct gct gtt aaa gaa tct tct aat gct 750  
 Asn Ala Ala Val Lys Glu Ser Ser Asn Ala 250  
 245

<210> 43  
 <211> 250  
 <212> PRT  
 <213> Naturally occurring gamma proteobacterium



<400> 43

Met Gly Lys Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr  
1 5 10 15

Phe Ala Ala Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val  
20 25 30

Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe  
35 40 45

Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
50 55 60

Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met  
65 70 75 80

Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr  
85 90 95

Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu  
100 105 110

Ile Leu Ala Ala Thr Asn Val Ala Gly Ser Leu Phe Lys Lys Leu  
115 120 125

Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala

09847513.000001

130	135	140	
Gly Ile Met Ala Ala Trp Pro Ala Phe Ile Ile Gly Cys Leu Ala Trp			
145	150	155	160
Val Tyr Met Ile Tyr Glu Leu Tyr Ala Gly Glu Gly Lys Ser Ala Cys			
	165	170	175
Asn Thr Ala Ser Pro Ser Val Gln Ser Ala Tyr Asn Thr Met Met Ala			
	180	185	190
Ile Ile Val Phe Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly			
	195	200	205
Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Ile Tyr			
	210	215	220
Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile Trp			
225	230	235	240
Asn Ala Ala Val Lys Glu Ser Ser Asn Ala			
	245	250	
<210>	44		
<211>	750		
<212>	DNA		
<213>	Naturally occurring gamma proteobacterium		

<220>  
 <221> CDS  
 <222> (1)..(750)  
 <223> Proteorhodopsin variant from pcr clone MB100m9; GenBank #AF349993

<400> 44 48  
 atg ggt aaa tta tta ctg ata tta ggt agt gtt att gca ctt cct aca  
 Met Gly Lys Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr  
 1 5 10 15  
 ttt gct gca ggt ggt ggt gac ctt gat gct agt gat tac act ggt gtt  
 Phe Ala Ala Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val  
 20 25 30 36  
 tct ttt tgg tta gtt act gct gct tta tta gca tct act gta ttt ttc  
 Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe  
 35 40 45 144  
 ttt gtt gaa aga gat aga gtt tct gca aaa tgg aaa aca tca tta act  
 Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
 50 55 60 192  
 gta tct ggt ctt gtt act ggt att gct ttc tgg cat tac atg tat atg  
 Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met  
 65 70 75 80 240  
 aga ggg gta tgg att gaa act ggt gat tcg cca act gta ttt aga tac  
 Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr  
 85 90 95 288  
 ata gat tgg tta cta cta gtt cct tta tta ata tgt gaa ttc tac tta  
 Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu  
 100 105 110 336

att ctt gcc gct gca act aat gtt gct gga tca tta ttt aag aaa tta 384  
 Ile Leu Ala Ala Thr Asn Val Ala Gly Ser Leu Phe Lys Lys Leu  
 115 120 125  
  
 ctt gtt ggt tct ctt gtt att gtt gtt ttt ggt tac atg ggt gaa gca 432  
 Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala  
 130 135 140  
  
 gga atc atg gct gca tgg cct gca ttc att att ggg tgt tta gct tgg 480  
 Gly Ile Met Ala Ala Trp Pro Ala Phe Ile Ile Gly Cys Leu Ala Trp  
 145 150 155 160  
  
 gta tac atg att tat gaa cta tgg gct gga gaa gga aaa tct gca tgt 528  
 Val Tyr Met Ile Tyr Glu Leu Trp Ala Gly Glu Lys Ser Ala Cys  
 165 170 175  
  
 aat act gca agt cct gct gct caa tca gct tac aac aca atg atg tat 576  
 Asn Thr Ala Ser Pro Ala Val Gln Ser Ala Tyr Asn Thr Met Met Tyr  
 180 185 190  
  
 atc atc atc ttt ggt tgg gcg att tat cct gta ggt tat ttc aca ggt 624  
 Ile Ile Ile Phe Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly  
 195 200 205  
  
 tac ctt atg ggt gac ggt gga tca gca ctt aac tta aac ctt att tat 672  
 Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Asn Leu Ile Tyr  
 210 215 220  
  
 aac ctt gct gac ttt gtt aac aag att cta ttt ggt tta att ata tgg 720  
 Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile Trp  
 225 230 235 240  
  
 aat gtt gct gtt aaa gaa tct tct aat gct  
 Asn Val Ala Val Lys Glu Ser Ser Asn Ala 750

09847513.0000001

245

250

<210> 45  
 <211> 250  
 <212> PRT  
 <213> Naturally occurring gamma proteobacterium

<400> 45

Met Gly Lys Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr  
 1 5 10 15

Phe Ala Ala Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val  
 20 25 30

Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe  
 35 40 45

Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
 50 55 60

Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met  
 65 70 75 80

Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr  
 85 90 95

Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu

09847513.0000001



Asn Val Ala Val Lys Glu Ser Ser Asn Ala  
245 250

<210> 46  
<211> 750  
<212> DNA  
<213> Naturally occurring gamma proteobacterium

<220>  
<221> CDS  
<222> (1)..(750)  
<223> Proteorhodopsin variant from pcr clone MB100m10; GenBank #AF34999

<400> 46  
atg ggt aaa tta tta ctg ata tta ggt agt gtt att gca ctt cct aca  
Met Gly Lys Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr  
1 5 10 15 48

ttt gct gca ggt ggc ggt gac ctt gat gct agt gat tac act ggt gtt  
Phe Ala Ala Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val  
20 25 30 96

tct ttt tgg tta gtt aca gct gct cta tta gcg tct act gta ttt ttc  
Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe  
35 40 45 144

ttt gtt gaa aga gat aga gtt tct gca aaa tgg aaa aca tca tta act  
Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
50 55 60 192

gta tct ggt ctt gtt act ggt att gct ttc tgg cat tac atg tat atg  
Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met  
240

65	70	75	80	
aga gga gta tgg att gaa act ggt gat tcg cca act gta ttt aga tac				288
Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr	85	90	95	
att gat tgg tta cta aca gtt cct tta tta ata tgt gaa ttc tac tta				336
Ile Asp Trp Leu Leu Thr Val Pro Leu Ile Cys Glu Phe Tyr Leu	100	105	110	
att ctt gct gct gca act aat gtt gcc ggc tca tta ttt aag aaa ctt				384
Ile Leu Ala Ala Thr Asn Val Ala Gly Ser Leu Phe Lys Lys Leu	115	120	125	
cta gtt ggt tct ctt gtt atg ctt gtg ttt ggt tac atg ggt gaa gca				432
Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala	130	135	140	
gga ata atg gcg gct tgg cct gca ttc atc gtt gga tgt tta gca tgg				480
Gly Ile Met Ala Ala Trp Pro Ala Phe Ile Val Gly Cys Leu Ala Trp	145	150	155	
gta tat atg att tat gaa cta tgg gct ggt gaa gga aaa tct gca tgt				528
Val Tyr Met Ile Tyr Glu Leu Trp Ala Gly Glu Gly Lys Ser Ala Cys	165	170	175	
aat act gca agt cct gct gta cag tca gct tac aac aca atg atg tat				576
Asn Thr Ala Ser Pro Ala Val Gln Ser Ala Tyr Asn Thr Met Met Tyr	180	185	190	
atc atc atc gtt ggt tgg gca att tat cct gta ggt tat ttc aca ggt				624
Ile Ile Ile Val Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly	195	200	205	
tac cta atg ggt gac ggt gga tca gct ctt aat cta aac ctt att tat				672

09847513, 0000001



Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Ile Tyr  
210 215 220

aac ctt gct gac ttt gtt aac aag att cta ttt ggt tta att ata tgg 720  
Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile Trp 240  
225 230 235

aat gtt gct gtt aaa gaa tct tct aat gct 750  
Asn Val Ala Val Lys Glu Ser Ser Asn Ala 250  
245

<210> 47  
<211> 250  
<212> PRT  
<213> Naturally occurring gamma proteobacterium

<400> 47

Met Gly Lys Leu Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr  
1 5 10 15

Phe Ala Ala Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val  
20 25 30

Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe  
35 40 45

Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
50 55 60

Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met  
 65 70 75 80  
 Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr  
 85 90 95  
 Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu  
 100 105 110  
 Ile Leu Ala Ala Thr Asn Val Ala Gly Ser Leu Phe Lys Lys Leu  
 115 120 125  
 Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala  
 130 135 140  
 Gly Ile Met Ala Ala Trp Pro Ala Phe Ile Val Gly Cys Leu Ala Trp  
 145 150 155 160  
 Val Tyr Met Ile Tyr Glu Leu Trp Ala Gly Glu Gly Lys Ser Ala Cys  
 165 170 175  
 Asn Thr Ala Ser Pro Ala Val Gln Ser Ala Tyr Asn Thr Met Met Tyr  
 180 185 190  
 Ile Ile Ile Val Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly  
 195 200 205

Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Asn Leu Ile Tyr  
 210 215 220

Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile Trp  
 225 230 235 240

Asn Val Ala Val Lys Glu Ser Ser Asn Ala  
 245 250

<210> 48  
 <211> 753  
 <212> DNA  
 <213> Naturally occurring gamma proteobacterium

<220>  
 <221> CDS  
 <222> (1)..(753)  
 <223> Proteorhodopsin variant from pcr clone PALB1; GenBank #AF349995

<400> 48  
 atg ggt aaa tta tta ctg ata tta ggt agt gct att gca ctt cca tca  
 Met Gly Lys Leu Leu Ile Leu Gly Ser Ala Ile Ala Leu Pro Ser  
 1 5 10 15

ttt gct gct gct ggt ggc gat cta gat ata agt gat act gtt ggt gtt  
 Phe Ala Ala Ala Gly Gly Asp Leu Asp Ile Ser Asp Thr Val Gly Val  
 20 25 30 96

tca ttc tgg ctg gtt aca gct ggt atg tta gcg gca act gta ttc ttt  
 Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Ala Thr Val Phe Phe  
 144

35	40	45	
ttt gta gaa aga gac caa gtc agc gct aag tgg aaa act tca ctt act			192
Phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Thr	55	60	
50			
gta tct ggt tta att act ggt ata gct ttt tgg cat tat ctc tac atg			240
Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met	70	75	80
65			
aga ggt gtt tgg ata gat act ggt gat aca cca aca gta ttt aga tat			288
Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr	85	90	95
att gat tgg cta tta act gtt cca tta caa atg gtt gag ttc tat cta			336
Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Met Val Glu Phe Tyr Leu	100	105	110
att ctt gct gct tgt tga aca agt gtt gct gct tca tta ttt aag aag ctt			384
Ile Leu Ala Ala Cys Thr Ser Val Ala Ala Ser Leu Phe Lys Lys Leu	115	120	125
cta gct ggt tca tta gta atg tta ggt gct gga ttt gca ggc gaa gct			432
Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Phe Ala Gly Glu Ala	130	135	140
ggt tta gct cct gta tta cct gct ttc att ctt ggt atg gct ggt tgg			480
Gly Leu Ala Pro Val Leu Pro Ala Phe Ile Leu Gly Met Ala Gly Trp	145	150	155
tta tac atg att tat gag cta cat atg ggt gaa ggt aag gct gct gta			528
Leu Tyr Met Ile Tyr Glu Leu His Met Gly Glu Lys Ala Ala Val	165	170	175
agt act gca agt cct gct gtt aac tct gct tac aat gca atg atg aag			576

Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Lys  
 180 185 190  
 att att gtt att gga tgg gca att tat cct gct gga tat gct gct ggt 624  
 Ile Ile Val Ile Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly  
 195 200 205  
 tac cta atg agt ggt gac ggt gta tac gct tca aac tta aac ctt ata 672  
 Tyr Leu Met Ser Gly Asp Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile  
 210 215 220  
 tat aac ctt gct gac ttt gtt aac aag att cta ttt ggt ttg atc att 720  
 Tyr Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile  
 225 230 235 240  
 tgg aat gtt gct gtt aaa gaa tct tct aat gct 753  
 Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala 250  
 245  
 <210> 49  
 <211> 251  
 <212> PRT  
 <213> Naturally occurring gamma proteobacterium  
 <400> 49  
 Met Gly Lys Leu Leu Ile Leu Gly Ser Ala Ile Ala Leu Pro Ser  
 1 5 10 15  
 Phe Ala Ala Ala Gly Gly Asp Leu Asp Ile Ser Asp Thr Val Gly Val  
 20 25 30

Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Ala Thr Val Phe Phe  
 35 40 45  
 Phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
 50 55 60  
 Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met  
 65 70 75 80  
 Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr  
 85 90 95  
 Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Met Val Glu Phe Tyr Leu  
 100 105 110  
 Ile Leu Ala Ala Cys Thr Ser Val Ala Ala Ser Leu Phe Lys Lys Leu  
 115 120 125  
 Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Phe Ala Gly Glu Ala  
 130 135 140  
 Gly Leu Ala Pro Val Leu Pro Ala Phe Ile Leu Gly Met Ala Gly Trp  
 145 150 155 160  
 Leu Tyr Met Ile Tyr Glu Leu His Met Gly Glu Lys Ala Ala Val  
 165 170 175

09/847,513 DeLong et al.

Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Lys  
180 185 190

Ile Ile Val Ile Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly  
195 200 205

Tyr Leu Met Ser Gly Asp Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile  
210 215 220

Tyr Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile  
225 230 235 240

Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala  
245 250

<210> 50

<211> 753

<212> DNA

<213> Naturally occurring gamma proteobacterium

<220>

<221> CDS

<222> (1)..(753)

<223> Proteorhodopsin variant from pcr clone PALB2; GenBank #AF349996

<400> 50

atg ggt aaa tta tta ctg ata tta ggt agt gct att gca ctt cca tca  
Met Gly Lys Leu Leu Leu Ile Leu Gly Ser Ala Ile Ala Leu Pro Ser

48

1                    5                    10                    15  
ttt gct gct gct ggt ggc gat cta gat ata agt gat act gtt ggt gtt                    96  
Phe Ala Ala Gly Gly Asp Leu Asp Ile Ser Asp Thr Val Gly Val  
20                    25                    30  
tca ttc tgg ctg gtt aca gct ggt atg tta gcg gca act gtg ttc ttt                    144  
Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Thr Val Phe  
35                    40                    45  
ttt gta gaa aga gac caa gtc agc gct gag tgg aaa act tca ctt act                    192  
Phe Val Glu Arg Asp Gln Val Ser Ala Glu Trp Lys Thr Ser Leu Thr  
50                    55                    60  
gta tct ggt tta att act ggt ata gct ttt tgg cat tat ctc tat atg                    240  
Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met  
65                    70                    75                    80  
aga ggt gtt tgg ata gat act ggt gat acc cca aca gta ttc aga tat                    288  
Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr  
85                    90                    95  
att gat tgg tta tta act gtt cca tta caa atg gtt gag ttc tat cta                    336  
Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Met Val Glu Phe Tyr Leu  
100                    105                    110  
att ctt gct gct tgt aca agt gtt gct gct tca tta ttt aag aag ctt                    384  
Ile Leu Ala Ala Cys Thr Ser Val Ala Ala Ser Leu Phe Lys Lys Leu  
115                    120                    125  
cta gct ggt tca tta gta atg tta ggt gct gga ttt gca ggc gaa gct                    432  
Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Phe Ala Gly Glu Ala  
130                    135                    140  
gga tta gct cct gta tta cct gct ttc att att ggt atg gct gga tgg                    480

09847513.0000001



Gly Leu Ala Pro Val Leu Pro Ala Phe Ile Ile Gly Met Ala Gly Trp  
 145 150 155 160  
 tta tac atg att tat gag cta tat atg ggt gaa ggt aag gct gct gta  
 Leu Tyr Met Ile Tyr Glu Leu Tyr Met Gly Glu Gly Lys Ala Ala Val  
 165 170 175 528  
 agt act gca agt cct gct gtt aac tct gca tac aac gca atg atg atg  
 Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Met  
 180 185 190 576  
 att att gtt gtt gga tgg gca att tat cct gct gga tat gct gct ggt  
 Ile Ile Val Val Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly  
 195 200 205 624  
 tac cta atg ggt ggc gaa ggt gta tac gct tca aac tta aac ctt ata  
 Tyr Leu Met Gly Gly Glu Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile  
 210 215 220 672  
 tat aac ctt gct gac ttt gtt aac aag att cta ttt ggt ttg atc att  
 Tyr Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile  
 225 230 235 720  
 tgg aat gtt gct gtt aaa gaa tct tct aat gct  
 Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala  
 245 250 753  
 <210> 51  
 <211> 251  
 <212> PRT  
 <213> Naturally occurring gamma proteobacterium  
 <400> 51

Met Gly Lys Leu Leu Ile Leu Gly Ser Ala Ile Ala Leu Pro Ser  
 1 5 10 15  
 Phe Ala Ala Gly Gly Asp Leu Asp Ile Ser Asp Thr Val Gly Val  
 20 25 30  
 Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Thr Val Phe  
 35 40 45  
 Phe Val Glu Arg Asp Gln Val Ser Ala Glu Trp Lys Thr Ser Leu Thr  
 50 55 60  
 Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met  
 65 70 75 80  
 Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr  
 85 90 95  
 Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Met Val Glu Phe Tyr Leu  
 100 105 110  
 Ile Leu Ala Ala Cys Thr Ser Val Ala Ala Ser Leu Phe Lys Lys Leu  
 115 120 125  
 Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Phe Ala Gly Glu Ala  
 130 135 140

09847513.0000001

Gly Leu Ala Pro Val Leu Pro Ala Phe Ile Ile Gly Met Ala Gly Trp  
145 150 155 160

Leu Tyr Met Ile Tyr Glu Leu Tyr Met Gly Glu Gly Lys Ala Ala Val  
165 170 175

Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Met  
180 185 190

Ile Ile Val Val Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly  
195 200 205

Tyr Leu Met Gly Gly Glu Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile  
210 215 220

Tyr Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile  
225 230 235 240

Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala  
245 250

<210> 52

<211> 753

<212> DNA

<213> Naturally occurring gamma proteobacterium

<220>

09847513.0000001

<221> CDS  
 <222> (1)..(753)  
 <223> Proteorhodopsin variant from pcr clone PALB5; GenBank#AF349997

```

<400> 52                                48
atg ggt aaa tta tta ctg ata tta ggt agt gct att gca ctt cca tca
Met Gly Lys Leu Leu Ile Leu Gly Ser Ala Ile Ala Leu Pro Ser
1      5      10      15

ttt gct gct gct ggc gat cta gat ata agt gat act gtt ggt gtt
Phe Ala Ala Gly Gly Asp Leu Asp Ile Ser Asp Thr Val Gly Val
20      25      30      35      40      45      50      55      60      65

tca ttc tgg ctg gtt aca gct ggt atg tta gcg gca act gtg ttc ttt
Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Thr Val Phe
35      40      45      50      55      60      65      70      75      80      85      90      95      100      105      110      115      120      125      130      135      140      145      150      155      160      165      170      175      180      185      190      195      200      205      210      215      220      225      230      235      240      245      250      255      260      265      270      275      280      285      290      295      300      305      310      315      320      325      330      335      340      345      350      355      360      365      370      375      380      385      390      395      400      405      410      415      420      425      430      435      440      445      450      455      460      465      470      475      480      485      490      495      500      505      510      515      520      525      530      535      540      545      550      555      560      565      570      575      580      585      590      595      600      605      610      615      620      625      630      635      640      645      650      655      660      665      670      675      680      685      690      695      700      705      710      715      720      725      730      735      740      745      750      755      760      765      770      775      780      785      790      795      800      805      810      815      820      825      830      835      840      845      850      855      860      865      870      875      880      885      890      895      900      905      910      915      920      925      930      935      940      945      950      955      960      965      970      975      980      985      990      995      1000      1005      1010      1015      1020      1025      1030      1035      1040      1045      1050      1055      1060      1065      1070      1075      1080      1085      1090      1095      1100      1105      1110      1115      1120      1125      1130      1135      1140      1145      1150      1155      1160      1165      1170      1175      1180      1185      1190      1195      1200      1205      1210      1215      1220      1225      1230      1235      1240      1245      1250      1255      1260      1265      1270      1275      1280      1285      1290      1295      1300      1305      1310      1315      1320      1325      1330      1335      1340      1345      1350      1355      1360      1365      1370      1375      1380      1385      1390      1395      1400      1405      1410      1415      1420      1425      1430      1435      1440      1445      1450      1455      1460      1465      1470      1475      1480      1485      1490      1495      1500      1505      1510      1515      1520      1525      1530      1535      1540      1545      1550      1555      1560      1565      1570      1575      1580      1585      1590      1595      1600      1605      1610      1615      1620      1625      1630      1635      1640      1645      1650      1655      1660      1665      1670      1675      1680      1685      1690      1695      1700      1705      1710      1715      1720      1725      1730      1735      1740      1745      1750      1755      1760      1765      1770      1775      1780      1785      1790      1795      1800      1805      1810      1815      1820      1825      1830      1835      1840      1845      1850      1855      1860      1865      1870      1875      1880      1885      1890      1895      1900      1905      1910      1915      1920      1925      1930      1935      1940      1945      1950      1955      1960      1965      1970      1975      1980      1985      1990      1995      2000      2005      2010      2015      2020      2025      2030      2035      2040      2045      2050      2055      2060      2065      2070      2075      2080      2085      2090      2095      2100      2105      2110      2115      2120      2125      2130      2135      2140      2145      2150      2155      2160      2165      2170      2175      2180      2185      2190      2195      2200      2205      2210      2215      2220      2225      2230      2235      2240      2245      2250      2255      2260      2265      2270      2275      2280      2285      2290      2295      2300      2305      2310      2315      2320      2325      2330      2335      2340      2345      2350      2355      2360      2365      2370      2375      2380      2385      2390      2395      2400      2405      2410      2415      2420      2425      2430      2435      2440      2445      2450      2455      2460      2465      2470      2475      2480      2485      2490      2495      2500      2505      2510      2515      2520      2525      2530      2535      2540      2545      2550      2555      2560      2565      2570      2575      2580      2585      2590      2595      2600      2605      2610      2615      2620      2625      2630      2635      2640      2645      2650      2655      2660      2665      2670      2675      2680      2685      2690      2695      2700      2705      2710      2715      2720      2725      2730      2735      2740      2745      2750      2755      2760      2765      2770      2775      2780      2785      2790      2795      2800      2805      2810      2815      2820      2825      2830      2835      2840      2845      2850      2855      2860      2865      2870      2875      2880      2885      2890      2895      2900      2905      2910      2915      2920      2925      2930      2935      2940      2945      2950      2955      2960      2965      2970      2975      2980      2985      2990      2995      3000      3005      3010      3015      3020      3025      3030      3035      3040      3045      3050      3055      3060      3065      3070      3075      3080      3085      3090      3095      3100      3105      3110      3115      3120      3125      3130      3135      3140      3145      3150      3155      3160      3165      3170      3175      3180      3185      3190      3195      3200      3205      3210      3215      3220      3225      3230      3235      3240      3245      3250      3255      3260      3265      3270      3275      3280      3285      3290      3295      3300      3305      3310      3315      3320      3325      3330      3335      3340      3345      3350      3355      3360      3365      3370      3375      3380      3385      3390      3395      3400      3405      3410      3415      3420      3425      3430      3435      3440      3445      3450      3455      3460      3465      3470      3475      3480      3485      3490      3495      3500      3505      3510      3515      3520      3525      3530      3535      3540      3545      3550      3555      3560      3565      3570      3575      3580      3585      3590      3595      3600      3605      3610      3615      3620      3625      3630      3635      3640      3645      3650      3655      3660      3665      3670      3675      3680      3685      3690      3695      3700      3705      3710      3715      3720      3725      3730      3735      3740      3745      3750      3755      3760      3765      3770      3775      3780      3785      3790      3795      3800      3805      3810      3815      3820      3825      3830      3835      3840      3845      3850      3855      3860      3865      3870      3875      3880      3885      3890      3895      3900      3905      3910      3915      3920      3925      3930      3935      3940      3945      3950      3955      3960      3965      3970      3975      3980      3985      3990      3995      4000      4005      4010      4015      4020      4025      4030      4035      4040      4045      4050      4055      4060      4065      4070      4075      4080      4085      4090      4095      4100      4105      4110      4115      4120      4125      4130      4135      4140      4145      4150      4155      4160      4165      4170      4175      4180      4185      4190      4195      4200      4205      4210      4215      4220      4225      4230      4235      4240      4245      4250      4255      4260      4265      4270      4275      4280      4285      4290      4295      4300      4305      4310      4315      4320      4325      4330      4335      4340      4345      4350      4355      4360      4365      4370      4375      4380      4385      4390      4395      4400      4405      4410      4415      4420      4425      4430      4435      4440      4445      4450      4455      4460      4465      4470      4475      4480      4485      4490      4495      4500      4505      4510      4515      4520      4525      4530      4535      4540      4545      4550      4555      4560      4565      4570      4575      4580      4585      4590      4595      4600      4605      4610      4615      4620      4625      4630      4635      4640      4645      4650      4655      4660      4665      4670      4675      4680      4685      4690      4695      4700      4705      4710      4715      4720      4725      4730      4735      4740      4745      4750      4755      4760      4765      4770      4775      4780      4785      4790      4795      4800      4805      4810      4815      4820      4825      4830      4835      4840      4845      4850      4855      4860      4865      4870      4875      4880      4885      4890      4895      4900      4905      4910      4915      4920      4925      4930      4935      4940      4945      4950      4955      4960      4965      4970      4975      4980      4985      4990      4995      5000      5005      5010      5015      5020      5025      5030      5035      5040      5045      5050      5055      5060      5065      5070      5075      5080      5085      5090      5095      5100      5105      5110      5115      5120      5125      5130      5135      5140      5145      5150      5155      5160      5165      5170      5175      5180      5185      5190      5195      5200      5205      5210      5215      5220      5225      5230      5235      5240      5245      5250      5255      5260      5265      5270      5275      5280      5285      5290      5295      5300      5305      5310      5315      5320      5325      5330      5335      5340      5345      5350      5355      5360      5365      5370      5375      5380      5385      5390      5395      5400      5405      5410      5415      5420      5425      5430      5435      5440      5445      5450      5455      5460      5465      5470      5475      5480      5485      5490      5495      5500      5505      5510      5515      5520      5525      5530      5535      5540      5545      5550      5555      5560      5565      5570      5575      5580      5585      5590      5595      5600      5605      5610      5615      5620      5625      5630      5635      5640      5645      5650      5655      5660      5665      5670      5675      5680      5685      5690      5695      5700      5705      5710      5715      5720      5725      5730      5735      5740      5745      5750      5755      5760      5765      5770      5775      5780      5785      5790      5795      5800      5805      5810      5815      5820      5825      5830      5835      5840      5845      5850      5855      5860      5865      5870      5875      5880      5885      5890      5895      5900      5905      5910      5915      5920      5925      5930      5935      5940      5945      5950      5955      5960      5965      5970      5975      5980      5985      5990      5995      6000      6005      6010      6015      6020      6025      6030      6035      6040      6045      6050      6055      6060      6065      6070      6075      6080      6085      6090      6095      6100      6105      6110      6115      6120      6125      6130      6135      6140      6145      6150      6155      6160      6165      6170      6175      6180      6185      6190      6195      6200      6205      6210      6215      6220      6225      6230      6235      6240      6245      6250      6255      6260      6265      6270      6275      6280      6285      6290      6295      6300      6305      6310      6315      6320      6325      6330      6335      6340      6345      6350      6355      6360      6365      6370      6375      6380      6385      6390      6395      6400      6405      6410      6415      6420      6425      6430      6435      6440      6445      6450      6455      6460      6465      6470      6475      6480      6485      6490      6495      6500      6505      6510      6515      6520      6525      6530      6535      6540      6545      6550      6555      6560      6565      6570      6575      6580      6585      6590      6595      6600      6605      6610      6615      6620      6625      6630      6635      6640      6645      6650      6655      6660      6665      6670      6675      6680      6685      6690      6695      6700      6705      6710      6715      6720      6725      6730      6735      6740      6745      6750      6755      6760      6765      6770      6775      6780      6785      6790      6795      6800      6805      6810      6815      6820      6825      6830      6835      6840      6845      6850      6855      6860      6865      6870      6875      6880      6885      6890      6895      6900      6905      6910      6915      6920      6925      6930      6935      6940      6945      6950      6955      6960      6965      6970      6975      6980      6985      6990      6995      7000      7005      7010      7015      7020      7025      7030      7035      7040      7045      7050      7055      7060      7065      7070      7075      7080      7085      7090      7095      7100      7105      7110      7115      7120      7125      7130      7135      7140      7145      7150      7155      7160      7165      7170      7175      7180      7185      7190      7195      7200      7205      7210      7215      7220      7225      7230      7235      7240      7245      7250      7255      7260      7265      7270      7275      7280      7285      7290      7295      7300      7305      7310      7315      7320      7325      7330      7335      7340      7345      7350      7355      7360      7365      7370      7375      7380      7385      7390      7395      7400      7405      7410      7415      7420      7425      7430      7435      7440      7445      7450      7455      7460      7465      7470      7475      7480      7485      7490      7495      7500      7505      7510      7515      7520      7525      7530      7535      7540      7545      7550      7555      7560      7565      7570      7575      7580      7585      7590      7595      7600      7605      7610      7615      7620      7625      7630      7635      7640      7645      7650      7655      7660      7665      7670      7675      7680      7685      7690      7695      7700      7705      7710      7715      7720      7725      7730      7735      7740      7745      7750      7755      7760      7765      7770      7775      7780      7785      7790      7795      7800      7805      7810      7815      7820      7825      7830      7835      7840      7845      7850      7855      7860      7865      7870      7875      7880      7885      7890      7895      7900      7905      7910      7915      7920      7925      7930      7935      7940      7945      7950      7955      7960      7965      7970      7975      7980      7985      7990      7995      8000      8005      8010      8015      8020      8025      8030      8035      8040      8045      8050      8055      8060      8065      8070      8075      8080      8085      8090      8095      8100      8105      8110      8115      8120      8125      8130      8135      8140      8145      8150      8155      8160      8165      8170      8175      8180      8185      8190      8195      8200      8205      8210      8215      8220      8225      8230      8235      8240      8245      8250      8255      8260      8265      8270      8275      8280      8285      8290      8295      8300      8305      8310      8315      8320      8325      8330      8335      8340      8345      8350      8355      8360      8365      8370      8375      8380      8385      8390      8395      8400      8405      8410      8415      8420      8425      8430      8435      8440      8445      8450      8455      8460      8465      8470      8475      8480      8485      8490      8495      8500      8505      8510      8515      8520      8525      8530      8535      8540      8545      8550      8555      8560      8565      8570      8575      8580      8585      8590      8595      8600      8605      8610      8615      8620      8625      8630      8635      8640      8645      8650      8655      8660      8665      8670      8675      8680      8685      8690      8695      8700      8705      8710      8715      8720      8725      8730      8735      8740      8745      8750      8755      8760      8765      8770      8775      8780      8785      8790      8795      8800      8805      8810      8815      8820      8825      8830      8835      8840      8845      8850      8855      8860      8865      8870      8875      8880      8885      8890      8895      8900      8905      8910      8915      8920      8925      8930      8935      8940      8945      8950      8955      8960      8965      8970      8975      8980      8985      8990      8995      9000      9005      9010      9015      9020      9025      9030      9035      9040      9045      9050      9055      9060      9065      9070      9075      9080      9085      9090      9095      9100      9105      9110      9115      9120      9125      9130      9135      9140      9145      9150      9155      9160      9165      9170      9175      9180      9185      9190      9195      9200      9205      9210      9215      9220      9225      9230      9235      9240      9245      9250      9255      9260      9265      9270      9275      9280      9285      9290      9295      9300      9305      9310      9315      9320      9325      9330      9335      9340      9345      9350      9355      9360      9365      9370      9375      9380      9385      9390      9395      9400      9405      9410      9415      9420      9425      9430      9435      9440      9445      9450      9455      9460      9465      9470      9475      9480      9485      9490      9495      9500      9505      9510      9515      9520      9525      9530      9535      9540      9545      9550      9555      9560      9565      9570      9575      9580      9585      9590      9595      9600      9605      9610      9615      9620      9625      9630      9635      9640      9645      9650      9655      9660      9665      9670      9675      9680      9685      9690      9695      9700      9705      9710      9715      9720      9725      9730      9735      9740      9745      9750      9755      9760      9765      9770      9775      9780      9785      9790      9795      9800      9805      9810      9815      9820      9825      9830      9835      9840      9845      9850      9855      9860      9865      9870      9875      9880      9885      9890      9895      9900      9905      9910      9915      9920      9925      9930      9935      9940      9945      9950      9955      9960      9965      9970      9975      9980      9985      9990      9995      10000      10005      10010      10015      10020      10025      10030      10035      10040      10045      10050      10055      10060      10065      10070      10075      10080      10085      10090      10095      10100      10105      10110      10115      10120      10125      10130      10135      10140      10145      10150      10155      10160      10165      10170      10175      10180      10185      10190      10195      10200     
```

Ile Leu Ala Ala Cys Thr Asn Val Ala Ala Ser Leu Phe Lys Lys Leu  
 115 120 125  
 cta gct ggt tca tta gta atg tta ggt gct gga ttt gca ggc gaa gct  
 Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Phe Ala Gly Glu Ala  
 130 135 140 432  
 gga tta gct cct gta tgg cct gct ttc att att ggt atg gct gga tgg  
 Gly Leu Ala Pro Val Trp Pro Ala Phe Ile Ile Gly Met Ala Gly Trp  
 145 150 155 480  
 tta tac atg att tat gag cta tat atg ggt gaa ggt Gly Lys Ala Val  
 Leu Tyr Met Ile Tyr Glu Leu Tyr Met Gly Glu 170 175 528  
 agt act gca agt cct gct gct gtt aac tct gca tac aac gca atg atg atg  
 Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Met  
 180 185 190 576  
 att att gtt gtt gga tgg gca att tat cct gct gga tat gct gct ggt  
 Ile Ile Val Val Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly  
 195 200 205 624  
 tac cta atg ggt ggc gaa ggt gta tac gct tca aac cta aac ctt ata  
 Tyr Leu Met Gly Gly Glu Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile  
 210 215 220 672  
 tat aac ctt gct gac ttt gtt aac aag att cta ttt ggt ttg atc att  
 Tyr Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile  
 225 230 235 720  
 tgg aat gtt gct gtt aaa gaa tct tct aat gct  
 Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala  
 245 250 753

09847513.000001

<210> 53  
 <211> 251  
 <212> PRT  
 <213> Naturally occurring gamma proteobacterium  
 <400> 53

Met Gly Lys Leu Leu Ile Leu Gly Ser Ala Ile Ala Leu Pro Ser  
 1 5 10 15

Phe Ala Ala Gly Gly Asp Leu Asp Ile Ser Asp Thr Val Gly Val  
 20 25 30

Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Ala Thr Val Phe Phe  
 35 40 45

Phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
 50 55 60

Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met  
 65 70 75 80

Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr  
 85 90 95

Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Met Val Glu Phe Tyr Leu  
 100 105 110

09847513.0000001

Ile Leu Ala Ala Cys Thr Asn Val Ala Ala Ser Leu Phe Lys Lys Leu  
115 120 125

Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Phe Ala Gly Glu Ala  
130 135 140

Gly Leu Ala Pro Val Trp Pro Ala Phe Ile Ile Gly Met Ala Gly Trp  
145 150 155 160

Leu Tyr Met Ile Tyr Glu Leu Tyr Met Gly Glu Gly Lys Ala Ala Val  
165 170 175

Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Met  
180 185 190

Ile Ile Val Val Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly  
195 200 205

Tyr Leu Met Gly Gly Glu Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile  
210 215 220

Tyr Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile  
225 230 235 240

Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala  
245 250

09847513.000001

<210> 54  
 <211> 753  
 <212> DNA  
 <213> Naturally occurring gamma proteobacterium  
  
 <220>  
 <221> CDS  
 <222> (1)..(753)  
 <223> Proteorhodopsin variant from pcr clone PalB7; GenBank #AF349999  
  
 <400> 54 48  
 atg ggt aaa tta tta ctg ata tta ggt agt gct att gcg ctt cca tca  
 Met Gly Lys Leu Leu Ile Leu Gly Ser Ala Ile Ala Leu Pro Ser  
 1 5 10 15  
  
 ttt gct gct gct ggt ggc gat cta gat ata agt gat act gtt ggt gtt 96  
 Phe Ala Ala Ala Gly Gly Asp Leu Asp Ile Ser Asp Thr Val Gly Val  
 20 25 30  
  
 tca ttc tgg ctg gtt acg gct gct ggt atg tta gcg gca act gta ttc ttt 144  
 Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Ala Thr Val Phe Phe  
 35 40 45  
  
 ttt gta gaa aga gac caa gtc agc gct aag tgg aaa act tca ctt act 192  
 Phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
 50 55 60  
  
 gta tct ggt tta att act ggt ata gct ttt tgg cat tat ctc tac atg 240  
 Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met  
 65 70 75 80  
  
 aga ggt gtt tgg ata gat act ggt gat aca cca aca gta ttt aga tat 288



Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr  
 85 90 95  
 att gat tgg tta tta act gtt cca tta caa atg gtt gag ttc tat cta 336  
 Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Met Val Glu Phe Tyr Leu  
 100 105 110  
 att ctt gcc gct tgt aca agt gtt gct gct tca tta ttt aag aag ctt 384  
 Ile Leu Ala Ala Cys Thr Ser Val Ala Ala Ser Leu Phe Lys Lys Leu  
 115 120 125  
 cta gct ggt tca ttg gta atg tta ggt gct gga tct gca ggc gaa gct 432  
 Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Ser Ala Gly Glu Ala  
 130 135 140  
 gga tta gct cct gta tta cct gct ttc att att ggt atg gct gga tgg 480  
 Gly Leu Ala Pro Val Leu Pro Ala Phe Ile Ile Gly Met Ala Gly Trp  
 145 150 155 160  
 tta tac atg att tat gag cta tat atg ggt gaa ggt aag gct gct gta 528  
 Leu Tyr Met Ile Tyr Glu Leu Tyr Met Gly Glu Gly Lys Ala Ala Val  
 165 170 175  
 agt act gca agt cct gct gtt aac tct gca tac aac gca atg atg atg 576  
 Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Met  
 180 185 190  
 att att gtt gtt gga tgg gca att tat cct gct gga tat gct gct ggt 624  
 Ile Ile Val Val Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly  
 195 200 205  
 tac cta atg ggt ggc gaa ggt gta tac gct tca aac tta aac ctc ata 672  
 Tyr Leu Met Gly Gly Glu Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile  
 210 215 220

tat aac ctt gct gac ttt gtt aac aag att cta ttt ggt ttg atc att 720  
 Tyr Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile 240  
 225 230 235

tgg aat gtt gct gtt aaa gaa tct tct aat gct 753  
 Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala 250  
 245

<210> 55  
 <211> 251  
 <212> PRT  
 <213> Naturally occurring gamma proteobacterium

<400> 55

Met Gly Lys Leu Leu Leu Ile Leu Gly Ser Ala Ile Ala Leu Pro Ser 15  
 1 5 10

Phe Ala Ala Ala Gly Gly Asp Leu Asp Ile Ser Asp Thr Val Gly Val 30  
 20 25

Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Ala Thr Val Phe Phe 45  
 35 40

Phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Thr 60  
 50 55

Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met 80  
 65 70 75

Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr  
 85 90 95  
 Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Met Val Glu Phe Tyr Leu  
 100 105 110  
 Ile Leu Ala Ala Cys Thr Ser Val Ala Ala Ser Leu Phe Lys Lys Leu  
 115 120 125  
 Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Ser Ala Gly Glu Ala  
 130 135 140  
 Gly Leu Ala Pro Val Leu Pro Ala Phe Ile Ile Gly Met Ala Gly Trp  
 145 150 155 160  
 Leu Tyr Met Ile Tyr Glu Leu Tyr Met Gly Glu Gly Lys Ala Ala Val  
 165 170 175  
 Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Met  
 180 185 190  
 Ile Ile Val Val Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly  
 195 200 205  
 Tyr Leu Met Gly Gly Glu Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile  
 210 215 220

Tyr Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile  
 225 230 235 240

Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala  
 245 250

<210> 56  
 <211> 753  
 <212> DNA  
 <213> Naturally occurring gamma proteobacterium

<220>  
 <221> CDS  
 <222> (1)..(753)  
 <223> Proteorhodopsin variant from pcr clone PalB6; GenBank # AF349998

<400> 56 48  
 atg ggt aaa tta tta ctg ata tta ggt agt gct att gca ctt cca tca  
 Met Gly Lys Leu Leu Ile Leu Gly Ser Ala Ile Ala Leu Pro Ser  
 5 10 15  
 1  
 ttt gct gct gct ggt ggc gat cta gat ata agt gat act gtt ggt gtt  
 Phe Ala Ala Ala Gly Asp Leu Asp Ile Ser Asp Thr Val Gly Val  
 20 25 30  
 96  
 tca ttc tgg ctg gtt aca gct ggt atg tta gcg gca act gtg ttc ttt  
 Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Ala Thr Val Phe Phe  
 35 40 45  
 144

192  
 ttt gta gaa aga gac caa gtc agc gct aag tgg aaa act tca ctt act  
 phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
 50 55 60  
  
 240  
 gta tct ggt tta att act ggt ata gct ttt tgg cat tat ctc tat atg  
 Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met  
 65 70 75 80  
  
 288  
 aga ggt gtt tgg ata gac act ggt gat acc cca aca gta ttc aga tat  
 Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr  
 85 90 95  
  
 336  
 att gat tgg tta tta act gtt cca tta caa atg gtt gag ttc tat cta  
 Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Met Val Glu Phe Tyr Leu  
 100 105 110  
  
 384  
 att ctt gct gct tgt aca aat gtt gct gct tca tta ttt aag aag ctt  
 Ile Leu Ala Ala Cys Thr Asn Val Ala Ala Ser Leu Phe Lys Lys Leu  
 115 120 125  
  
 432  
 cta gct ggt tca tta gta atg tta ggt gct gga ttt gca ggc gaa gct  
 Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Phe Ala Gly Glu Ala  
 130 135 140  
  
 480  
 gga tta gct cct gta tgg cct gct ttc att att ggt atg gct gga tgg  
 Gly Leu Ala Pro Val Trp Pro Ala Phe Ile Ile Gly Met Ala Gly Trp  
 145 150 155 160  
  
 528  
 tta tac atg att tat gag cta tat atg ggt gaa ggt aag gct gct gta  
 Leu Tyr Met Ile Tyr Glu Leu Tyr Met Gly Glu Gly Lys Ala Ala Val  
 165 170 175  
  
 576  
 agt act gca agt cct gct gtt aac tct gca tac aac gca atg atg gtg  
 Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Val  
 180 185 190

09/847,513 DeLong et al.

624

att att gtt gtt gga tgg gca att tat cct gct gga tat gct gct ggt  
Ile Ile Val Val Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly  
195 200 205

672

tac cta atg ggt ggc gaa ggt gta tac gct tca aac cta aac ctt ata  
Tyr Leu Met Gly Gly Glu Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile  
210 215 220

720

tat aac ctt gct gac ttt gtt aac aag att cta ttt ggt ttg atc att  
Tyr Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile  
225 230 235 240

753

tgg aat gtt gct gtt aaa gaa tct tct aat gct  
Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala  
245 250

<210> 57

<211> 251

<212> PRT

<213> Naturally occurring gamma proteobacterium

<400> 57

Met Gly Lys Leu Leu Leu Ile Leu Gly Ser Ala Ile Ala Leu Pro Ser  
1 5 10 15

Phe Ala Ala Ala Gly Gly Asp Leu Asp Ile Ser Asp Thr Val Gly Val  
20 25 30

Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Thr Val Phe Phe  
35 40 45

Phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
 50 55 60  
 Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met  
 65 70 75 80  
 Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr  
 85 90 95  
 Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Met Val Glu Phe Tyr Leu  
 100 105 110  
 Ile Leu Ala Ala Cys Thr Asn Val Ala Ala Ser Leu Phe Lys Lys Leu  
 115 120 125  
 Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Phe Ala Gly Glu Ala  
 130 135 140  
 Gly Leu Ala Pro Val Trp Pro Ala Phe Ile Ile Gly Met Ala Gly Trp  
 145 150 155 160  
 Leu Tyr Met Ile Tyr Glu Leu Tyr Met Gly Glu Gly Lys Ala Ala Val  
 165 170 175  
 Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Val

180

185

190

Ile Ile Val Val Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly  
195 200 205

Tyr Leu Met Gly Gly Glu Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile  
210 215 220

Tyr Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile  
225 230 235 240

Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala  
245 250

&lt;210&gt; 58

&lt;211&gt; 753

&lt;212&gt; DNA

&lt;213&gt; Naturally occurring gamma proteobacteria

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(753)

&lt;223&gt; Proteorhodopsin variant from pcr clone PalB8; GenBank #AF350000

&lt;400&gt; 58

atg ggt aaa tta tta ctg ata tta ggt agt gct att gca ctt cca tca 48  
Met Gly Lys Leu Leu Ile Leu Gly Ser Ala Ile Ala Leu Pro Ser  
1 5 10 15

00047513.0000001



ttt gct gct gct ggt ggc gat cta gat ata agt gat act gtt ggt gtt 96  
 Phe Ala Ala Ala Gly Gly Asp Leu Asp Ile Ser Asp Thr Val Gly Val  
 20 25 30  
 tca ttc tgg ctg gtt aca gct ggt atg tta gcg gca act gtg ttc ttt 144  
 Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Ala Thr Val Phe Phe  
 35 40 45  
 ttt gta gaa aga gac caa gtc agc gct aag tgg aaa act tca ctt act 192  
 Phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
 50 55 60  
 gta tct ggt tta att act ggt ata gct ttt tgg cat tat ctc tat atg 240  
 Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met  
 65 70 75 80  
 aga ggt gtt tgg ata gac act ggt gat acc cca aca gta ttc aga tat 288  
 Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr  
 85 90 95  
 att gat tgg tta tta act gtt cca tta caa atg gtt gag ttc tat cta 336  
 Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Met Val Glu Phe Tyr Leu  
 100 105 110  
 att ctt gct gct tgt aca agt gtt gct gct tca tta ttt aag aag ctt 384  
 Ile Leu Ala Ala Cys Thr Ser Val Ala Ala Ser Leu Phe Lys Lys Leu  
 115 120 125  
 cta gct ggt tca tta gta atg tta ggt gct gga ttt gca ggc gaa gct 432  
 Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Phe Ala Gly Glu Ala  
 130 135 140  
 gga tta gct cct gta tta cct gct ttc att att ggt atg gct gga tgg 480  
 Gly Leu Ala Pro Val Leu Pro Ala Phe Ile Ile Gly Met Ala Gly Trp  
 145 150 155 160

tta tac atg att tat gag cta tat atg ggt gaa ggt aag gct gct gta 528  
 Leu Tyr Met Ile Tyr Glu Leu Tyr Met Gly Glu Lys Ala Ala Val 175  
 165 170  
 agt act gca agt cct gct gtt aac tct gca tac aac gca atg atg atg 576  
 Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Met 180 185 190  
 att att gtt gtt gga tgg gca att tat cct gct gga tat gct gct ggt 624  
 Ile Ile Val Val Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly 195 200 205  
 tac cta atg ggt ggc gaa ggt gta tac gct tca aac tta aac ctt ata 672  
 Tyr Leu Met Gly Gly Glu Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile 210 215 220  
 tat aac ctt gct gac ctt gtt aac aag att cta ttt ggt ttg atc att 720  
 Tyr Asn Leu Ala Asp Leu Val Asn Lys Ile Leu Phe Gly Leu Ile Ile 225 230 235 240  
 tgg aat gtt gct gtt aaa gaa gaa tct tct aat gct 753  
 Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala 245 250  
 <210> 59  
 <211> 251  
 <212> PRT  
 <213> Naturally occurring gamma proteobacteria  
 <400> 59  
 Met Gly Lys Leu Leu Leu Ile Leu Gly Ser Ala Ile Ala Leu Pro Ser 15  
 1 5 10

Phe Ala Ala Ala Gly Asp Leu Asp Ile Ser Asp Thr Val Gly Val  
20 25 30

Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Ala Thr Val Phe  
35 40 45

Phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
50 55 60

Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met  
65 70 75 80

Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr  
85 90 95

Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Met Val Glu Phe Tyr Leu  
100 105 110

Ile Leu Ala Ala Cys Thr Ser Val Ala Ala Ser Leu Phe Lys Lys Leu  
115 120 125

Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Phe Ala Gly Glu Ala  
130 135 140

Gly Leu Ala Pro Val Leu Pro Ala Phe Ile Ile Gly Met Ala Gly Trp

145 150 155 160

Leu Tyr Met Ile Tyr Glu Leu Tyr Met Gly Glu Gly Lys Ala Ala Val  
165 170 175

Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Met  
180 185 190

Ile Ile Val Val Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly  
195 200 205

Tyr Leu Met Gly Gly Glu Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile  
210 215 220

Tyr Asn Leu Ala Asp Leu Val Asn Lys Ile Leu Phe Gly Leu Ile Ile  
225 230 235 240

Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala  
245 250

<210> 60  
<211> 753  
<212> DNA  
<213> Naturally occurring gamma proteobacteria

<220>  
<221> CDS  
<222> (1)..(753)

<223> Proteorhodopsin variant from pcr clone PalE1;GenBank# AF350001

48	<400>	60	atg ggt aaa tta tta ctg ata tta ggt agt gct att gca ctt cca tca Met Gly Lys Leu Leu Ile Leu Gly Ser Ala Ile Ala Leu Pro Ser 1 5 10 15
96	ttt gct gct gct ggt ggc gat cta gat ata agt gat act gtt ggt gtt Phe Ala Ala Ala Gly Asp Leu Asp Ile Ser Asp Thr Val Gly Val 20 25 30		
144	tca ttc tgg ctg gtt aca gct ggt atg tta gcg gca act gtg ttc ttt Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Ala Thr Val Phe 35 40 45		
192	ttt gta gaa aga gac caa gtc agc gct aag tgg aaa act tca ctt act Phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Thr 50 55 60		
240	gta tct ggt tta att act ggt ata gct ttt tgg cat tat ctc tat atg Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met 65 70 75 80		
288	aga ggt gtt tgg ata gac act ggt gat acc cca aca gta ttc aga tat Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr 85 90 95		
336	att gat tgg tta tta act gtt cca tta caa gtg gtt gag ttc tat cta Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Val Val Glu Phe Tyr Leu 100 105 110		
384	att ctt gct gct tgt aca agt gtt gct gct tca tta ttt aag aag ctt Ile Leu Ala Ala Cys Thr Ser Val Ala Ala Ser Leu Phe Lys Lys Leu 115 120 125		



<211> 251  
 <212> PRT  
 <213> Naturally occurring gamma proteobacteria  
  
 <400> 61  
 Met Gly Lys Leu Leu Ile Leu Gly Ser Ala Ile Ala Leu Pro Ser  
 1 5 10 15  
 Phe Ala Ala Ala Gly Gly Asp Leu Asp Ile Ser Asp Thr Val Gly Val  
 20 25 30  
 Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Ala Thr Val Phe Phe  
 35 40 45  
 Phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
 50 55 60  
 Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met  
 65 70 75 80  
 Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr  
 85 90 95  
 Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Val Val Glu Phe Tyr Leu  
 100 105 110  
 Ile Leu Ala Ala Cys Thr Ser Val Ala Ala Ser Leu Phe Lys Lys Leu

115	120	125
Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Phe Ala Gly Glu Ala		
130	135	140
Gly Leu Ala Pro Val Leu Pro Ala Phe Ile Ile Gly Met Ala Gly Trp		
145	150	155
165	170	175
Leu Tyr Met Ile Tyr Glu Leu Tyr Met Gly Glu Gly Lys Ala Ala Val		
180	185	190
Ser Thr Ala Ser Pro Ala Val Asn Pro Ala Tyr Asn Ala Met Met Met		
195	200	205
Ile Ile Val Val Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly		
210	215	220
Tyr Leu Met Gly Gly Glu Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile		
225	230	235
Tyr Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile		
245	250	
Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala		

09/847,513 DeLong et al.



<210> 62  
 <211> 753  
 <212> DNA  
 <213> Naturally occurring gamma proteobacterium  
  
 <220>  
 <221> CDS  
 <222> (1)..(753)  
 <223> Proteorhodopsin variant from pcr clone Pale6; GenBank#AF350002

<400> 62		48
atg ggt aaa tta tta ctg ata tta ggt agt gct att gca ctt cca tca		
Met Gly Lys Leu Leu Ile Leu Gly Ser Ala Ile Ala Leu Pro Ser	15	
1	10	
5		
ttt gct gct gct ggt ggc gat cta gat ata agt gat act gtt ggt gtt		96
Phe Ala Ala Ala Gly Gly Asp Leu Asp Ile Ser Asp Thr Val Gly Val	30	
20	25	
35	40	
tca ttc tgg ctg gtt aca gct ggt atg tta gcg gca act gta ttc ttt		144
Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Ala Thr Val Phe	45	
192		
ttt gta gaa aga gac caa gtc agc gct aag tgg aaa act tca ctt act		
Phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Thr	60	
50	55	
gta tct ggt tta att act ggt ata gct ttt tgg cat tat ctc tac atg		240
Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met	75	
65	70	
aga ggt gtt tgg ata gat act ggt gat aca cca gta ttt aga tat		288
Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr	90	
85	95	

336	att gat tgg tta tta act gtt cca tta caa atg gtt gag ttc tat cta Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Met Val Glu Phe Tyr Leu 100 105 110
384	att ctt gct gct tgt aca agt gtt gct gct tca tta ttt aag aag ctt Ile Leu Ala Ala Cys Thr Ser Val Ala Ala Ser Leu Phe Lys Lys Leu 115 120 125
432	cta gct ggt tca tta gta atg tta ggt gct gga ttt gca ggc gaa gct Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Phe Ala Gly Glu Ala 130 135 140
480	ggt tta gct cct gta tta cct gct ttc att att ggt atg gct gga tgg Gly Leu Ala Pro Val Leu Pro Ala Phe Ile Ile Gly Met Ala Gly Trp 145 150 155 160
528	tta tac atg att tat gag cta cat atg ggt gaa ggt atg gct gct gta Leu Tyr Met Ile Tyr Glu Leu His Met Gly Glu Gly Lys Ala Ala Val 165 170 175
576	agt act gca agt cct gct gtt aac tct gca tac aac gca atg atg aag Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Lys 180 185 190
624	att att gtt att gga tgg gca att tat cct gct gga tat gct gct ggt Ile Ile Val Ile Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly 195 200 205
672	tac cta atg agt ggt gac ggt gta tac gct tca aac tta aac ctt ata Tyr Leu Met Ser Gly Asp Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile 210 215 220
720	tat aac ctt gct gac ttt gtt aac aag att cta ttt ggt ttg atc att Tyr Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile 225 230 235

Corrected Sequence Listing (August 4<sup>th</sup>, 2001)

228/235

225	230	235	240	
tgg aat gtt gct gtt aaa gaa tct tct aat gct				753
Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala				
245				250
<210> 63				
<211> 251				
<212> PRT				
<213> Naturally occurring gamma proteobacterium				
<400> 63				
Met Gly Lys Leu Leu Leu Ile Leu Gly Ser Ala Ile Ala Leu Pro Ser				
1	5	10	15	
Phe Ala Ala Ala Gly Gly Asp Leu Asp Ile Ser Asp Thr Val Gly Val				
20	25	30		
Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Ala Thr Val Phe Phe				
35	40	45		
Phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Thr				
50	55	60		
Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met				
65	70	75	80	
Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr				

00047513.0000001

Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Met Val Glu Phe Tyr Leu	85	90	95
100	105	110	
Ile Leu Ala Ala Cys Thr Ser Val Ala Ala Ser Leu Phe Lys Lys Leu	115	120	125
130	135	140	
Gly Leu Ala Pro Val Leu Pro Ala Phe Ile Ile Gly Met Ala Gly Trp	145	150	155
160	165	170	175
Leu Tyr Met Ile Tyr Glu Leu His Met Gly Glu Gly Lys Ala Ala Val	180	185	190
Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Lys	195	200	205
Ile Ile Val Ile Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly	210	215	220
Tyr Leu Met Ser Gly Asp Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile	225	230	235

09/847,513, 0000001

Tyr Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile  
 225 230 235 240

Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala  
 245 250

<210> 64  
 <211> 753  
 <212> DNA  
 <213> Naturally occurring gamma proteobacterium

<220>  
 <221> CDS  
 <222> (1)..(753)  
 <223> Proteorhodopsin variant from pcr clone PalE7; GenBank# AF350003

<400> 64 48  
 atg ggt aaa tta tta ctg ata tta ggt agt gct att gca ctt cca tca  
 Met Gly Lys Leu Leu Ile Leu Gly Ser Ala Ile Ala Leu Pro Ser  
 1 5 10 15

ttt gct gct gct ggt ggc gat cta gat ata agt gat act gtt ggt gtt  
 Phe Ala Ala Ala Gly Asp Leu Asp Ile Ser Asp Thr Val Gly Val  
 20 25 30 96

tca ttc tgg ctg gtt aca gct ggt atg tta gcg gca act gtg ttc ttt  
 Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Ala Thr Val Phe Phe  
 35 40 45 144

ttt gta gaa aga gac caa gtc agc gct aag tgg aaa act tca ctt act  
 Phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Thr  
 50 55 60 192

gta tct ggt tta att act ggt ata gct ttt tgg cat tat ctc tat atg 240  
 Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met  
 65 70 75 80  
 aga ggt gtt tgg ata gat act ggt gat acc cca aca gta ttc aga tat 288  
 Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr  
 85 90 95  
 att gat tgg tta tta act gtt cca tta caa atg gtt gag ttc tat cta 336  
 Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Met Val Glu Phe Tyr Leu  
 100 105 110  
 att ctt gct gct tgt aca agt gtt gct gct tca tta ttt aag aag ctt 384  
 Ile Leu Ala Ala Cys Thr Ser Val Ala Ala Ser Leu Phe Lys Lys Leu  
 115 120 125  
 cta gct ggt tca tta gta atg tta ggt gct gga ttt gca ggc gaa gct 432  
 Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Phe Ala Gly Glu Ala  
 130 135 140  
 gga tta gct cct gta tta cct gct ttc att att ggt atg gct gga tgg 480  
 Gly Leu Ala Pro Val Leu Pro Ala Phe Ile Ile Gly Met Ala Gly Trp  
 145 150 155 160  
 cta tac atg att tat gag cta tat atg ggt gaa ggt aag gct gct gta 528  
 Leu Tyr Met Ile Tyr Glu Leu Tyr Met Gly Glu Gly Lys Ala Ala Val  
 165 170 175  
 agt act gca agt cct gct gtt aac tct gca tac aac gca atg atg atg 576  
 Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Met  
 180 185 190  
 att att gtt gtt gga tgg gca att tat cct gct gga tat gct gct ggt 624  
 Ile Ile Val Val Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly

195	200	205	
tac cta atg ggt ggc gaa ggc gta tac gct tca aac tta aac ctt ata			672
Tyr Leu Met Gly Gly Glu Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile	210	215	
		220	
tat aac ctt gct gac ttt gtt aac aag att cta ttt ggt ttg atc att			720
Tyr Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile	225	230	
	235	240	
tgg aat gtt gct gtt aaa gaa tct tct aat gct			753
Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala	245	250	
<210> 65			
<211> 251			
<212> PRT			
<213> Naturally occurring gamma proteobacterium			
<400> 65			
Met Gly Lys Leu Leu Ile Leu Gly Ser Ala Ile Ala Leu Pro Ser	5	10	15
1			
Phe Ala Ala Ala Gly Gly Asp Leu Asp Ile Ser Asp Thr Val Gly Val	20	25	30
Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Ala Thr Val Phe Phe	35	40	45
Phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Thr			

50                    55                    60  
 Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met  
 65                    70                    75                    80  
 Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr  
 85                    90                    95  
 Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Met Val Glu Phe Tyr Leu  
 100                    105                    110  
 Ile Leu Ala Ala Cys Thr Ser Val Ala Ala Ser Leu Phe Lys Lys Leu  
 115                    120                    125  
 Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Phe Ala Gly Glu Ala  
 130                    135                    140  
 Gly Leu Ala Pro Val Leu Pro Ala Phe Ile Ile Gly Met Ala Gly Trp  
 145                    150                    155                    160  
 Leu Tyr Met Ile Tyr Glu Leu Tyr Met Gly Glu Gly Lys Ala Ala Val  
 165                    170                    175  
 Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Met  
 180                    185                    190

00047513.000001



Ile Ile Val Val Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly  
195 200

Tyr Leu Met Gly Gly Glu Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile  
210 215

Tyr Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile  
225 230 235 240

Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala  
245 250